

MESTRADO EM ONCOLOGIA
ESPECIALIZAÇÃO EM ONCOLOGIA LABORATORIAL

The impact of glycans as regulators of T cell-mediated immune response in gastrointestinal disorders

Bruno Miguel Maio Oliveira

M
2021



Bruno Miguel Maio Oliveira. **The impact of glycans as regulators of T cell-mediated immune response in gastrointestinal disorders**



The impact of glycans as regulators of T cell-mediated immune response in gastrointestinal disorders

Bruno Miguel Maio Oliveira



Bruno Miguel Maio Oliveira

THE IMPACT OF GLYCANS AS REGULATORS OF T CELL-MEDIATED IMMUNE RESPONSE IN GASTROINTESTINAL DISORDERS

Dissertação de Candidatura ao grau de Mestre em Oncologia – Especialização em Oncologia Laboratorial, submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientadora: Prof. Doutora Salomé S. Pinho, DVM, PhD

Categoria - Investigadora Principal (PI) do grupo “*Immunology, Cancer & Glycomedicine*”; Professora afiliada

Afiliação – Instituto de Investigação e Inovação em Saúde (i3S); Instituto de Ciências Biomédicas Abel Salazar

Coorientadora: Doutora Ângela Fernandes, PhD

Categoria - Investigadora de pós-doutoramento

Afiliação - Instituto de Investigação e Inovação em Saúde (i3S)

Coorientadora: Doutora Ana Dias*, PhD

Categoria - Investigadora Júnior

Afiliação - Instituto de Investigação e Inovação em Saúde (i3S)

*Coorientadora do projeto no modelo de CAC desde 1 de setembro de 2020 até 1 de julho de 2021

Agradecimentos

Primeiramente, agradeço à Professora Doutora Salomé Pinho por ter aceitado que desenvolvesse este trabalho no seu grupo de investigação do i3S - Instituto de Investigação e Inovação em Saúde da Universidade do Porto. No primeiro dia em que a conheci, a Professora Doutora Salomé disse-me que este projeto seria um grande desafio e foram nessas suas palavras que encontrei certezas que o desenvolvimento da minha dissertação de mestrado teria de ser sob sua orientação. Findo este período, não poderia estar mais agradecido pela incrível oportunidade que me deu, por todo o conhecimento que me transmitiu e por todo o apoio que proporcionou.

À Doutora Ângela Fernandes, minha coorientadora, um enorme agradecimento por toda a paciência, todo o apoio e todos os ensinamentos. Todos os dias me incentivaste a mostrar o meu melhor, e sempre me ajudaste quando não o consegui fazer. Um especial obrigado por me motivares a querer sempre melhorar enquanto investigador e pessoa.

À Doutora Ana Dias, também minha coorientadora, agradeço toda a disponibilidade e apoio no desenrolar deste projecto. Tal como a Ângela, foste um pilar fundamental para que conseguisse concluir esta etapa com sucesso. Obrigado por todo o carinho e pela amizade!

A todos os membros do Immunology, Cancer and Glycomedicine Group, agradeço a forma como me receberam e me integraram na equipa. Obrigado por toda ajuda que me prestaram no laboratório, pela disponibilidade em me esclarecer dúvidas e por contribuírem para o meu crescimento pessoal. Um especial obrigado à Catarina e à Mariana por me ajudarem incansavelmente para que conseguisse cumprir todos os objetivos a que me comprometi.

À Professora Doutora Cármen Jerónimo e toda a equipa de coordenação do Mestrado em Oncologia, agradeço pela formação de excelência que proporcionaram e que foi fundamental para que estivesse preparado para realizar este trabalho.

Ao Pedro, um obrigado por me acompanhares ao longo desta etapa tão importante. Estiveste sempre ao meu lado, nesta e noutras conquistas. Sabes sempre como me motivar e tens sempre as palavras certas para me animar. Obrigado por seres a pessoa mais especial da minha vida!

À Patricia, à Rita, à Catarina Gonçalves, ao Valério e à Catarina Vareiro, um especial obrigado por serem os melhores amigos que poderia pedir. Obrigado pelo companheirismo,

por me animarem sempre que estava em baixo e por me darem confiança para enfrentar todos os obstáculos.

Ao Jorge, o meu padrinho de Licenciatura, um obrigado pela grande amizade de tantos anos. Obrigado pelo excelente exemplo que me dás, tanto como investigador como pessoa. Obrigado por me apoiares todos os dias, mesmo estando longe.

Ao Vasco e ao Filipe, os meus afilhados, obrigado pela amizade e todo o apoio. Um abraço com saudade para vocês!

À minha irmã, Rita, obrigado por me inspirares todos os dias com a tua força e perseverança. Todos os dias motivas-me a não desistir, mesmo quando já não tenho forças para continuar. Obrigado pelo carinho e por seres a melhor irmã do mundo.

Aos meus pais, Carlos e Claudia, obrigado por todos os sacrifícios que fizeram para que conseguisse sempre alcançar os meus sonhos. Esta conquista é tanto minha como vossa. Irei deixar-vos cada vez mais orgulhosos de mim!

À minha restante família, particularmente à minha avó Rita, à minha avó Rosa, à minha bisavó Deolinda e ao meu tio Carlos, um obrigado por todo o apoio incondicional ao longo do meu percurso académico.

A todas as pessoas que contribuíram diretamente ou indiretamente para que conseguisse concluir esta etapa, fica aqui o meu mais sincero obrigado!

Extended dissertation synopsis

Colorectal cancer (CRC) is a malignant process that arises in the colon and rectum and a major public health issue that has increased in incidence and mortality over the past few decades, placing significant importance in further understanding this disease. Patient survival is heavily dependent on early detection. Although improvements in screening methods and early detection have contributed to the decrease in CRC-associated mortality, several factors, such as patient adherence to screening programs, access to infrastructure, and access to quality healthcare resources, influence the rate of success of early diagnosis and prognosis. Remarkably, only around 40% of patients are screened with early-stage disease. Diagnosis with an advanced-stage cancer (such as metastatic CRC) significantly reduces patient survival to around 12%. This reduced survival is also associated with the lack of therapeutic efficiency. Currently, surgery, chemo, and radiotherapy are the standard procedures for patient treatment. Novel treatments, such as immunotherapies, are continuously being developed but are still not eligible to the majority of CRC patients and have reduced effectiveness in controlling the disease. Therefore, there is a great need to identify the mechanisms underlying CRC development aiming to discover novel therapies and improve existing ones in order to reduce the burden of CRC and improve patients quality of life.

CRC can occur sporadically or in familial clusters due to the transmission of mutated alleles to descendants (hereditary CRC). Furthermore, it is a highly heterogeneous pathology defined by distinct cumulative genetic and epigenetic alterations, denominated as CRC pathways. These pathways result in distinct cancer subtypes, which are molecularly and phenotypically distinct from one another. Particularly important in the scope of this work, CpG island methylator phenotype (CIMP) and microsatellite instability (MSI) are the two main CRC pathways implicated in serrated CRC. This subtype presents lesions with a characteristic “serrated” morphology and does not follow the traditional adenoma-carcinoma sequence. Instead, serrated CRC relies on the progression from non-dysplastic serrated adenomas to dysplastic serrated adenomas by methylation or loss of the *MLH1* gene, with a consequent rapid evolution to serrated adenocarcinoma. Meanwhile, MSI is responsible for Lynch syndrome, a hereditary form of CRC, and involves the loss of mismatch repair (MMR) proteins function. Serrated and Lynch syndrome CRC lesions are characterized by an intense immune infiltrate. Moreover, the detection of precursor lesions in these subtypes is difficult, compromising long-term patient survival. This highlights an urgent need to study premalignant lesions aiming to develop novel targets to prevent disease progression.

Inflammatory bowel disease (IBD) is a chronic inflammatory disorder characterized by persistent inflammation of the gastrointestinal tract, severely affecting patients' quality of life. IBD encompasses both ulcerative colitis (UC) and Crohn's disease (CD). There is currently no cure for IBD, and existing therapies lack effectiveness in achieving disease remission. Furthermore, IBD patients have an increased risk of developing malignancies in the intestine, designated as colitis-associated colorectal cancer (CAC), which is associated with disease extension and duration. Since the burden of IBD is increasing and current screening methods often fail to detect CAC in its early stages, we can expect that this will be accompanied by an increased incidence of CAC. For this reason, there is a great need to improve our knowledge regarding the colitis-associated cancer cascade in order to uncover solutions for this emerging problem.

The immune system is an important defense mechanism programmed to protect the host against any threat. Disturbance of the immune system homeostasis can lead to several diseases, namely autoimmune diseases, including IBD, and cancer. In fact, cancer cells employ a myriad of mechanisms that allow evasion from immune surveillance.

N-glycosylation is a frequent post-translational modification occurring in ~90% of glycoproteins. *N*-glycans arise from the secretory pathway, which forms three distinct structure types: high mannose, complex, and hybrid *N*-glycans. *N*-glycans are particularly important in T cell development and activity. Interestingly, *N*-glycosylation plays an important role in mediating the pathogenesis of IBD. T cells from these patients exhibit decreased levels of β -1,6-GlcNAc branched *N*-glycans on the T cell receptor (TCR), which is associated with a decreased activation threshold and hyperactive immune response. Furthermore, our group has also demonstrated that CRC cells are characterized by increased surface expression of β -1,6-GlcNAc branched *N*-glycans showing that these tumor-associated glycans structures act as an important immune checkpoint in cancer. Glycoreprogramming these cells to express less complex *N*-glycans potentiated an anti-cancer immune response and led to CRC suppression .

In this work, we aimed to uncover how β -1,6-GlcNAc branched *N*-glycans impact T cell activity along distinct CRC pathways: serrated pathway, Lynch syndrome, and CAC cascades. Using well-characterized sample cohorts from patients in these cascades (including paraffin samples and fresh biopsies), we found evidence suggesting that β -1,6-GlcNAc branched *N*-glycans are overexpressed in T cells in premalignant serrated and Lynch syndrome lesions, which in turn can be a factor responsible for further disease progression. Our preliminary results from the CAC cascade also suggest that these complex glycans structures are overexpressed in intermediate and high risk colitis, which seem to be associated with the risk of progression from colitis to dysplasia. We also observed an

increase in infiltration of regulatory T cells (Tregs) in low risk and intermediate risk colitis, suggesting that the T cells glyco profile might be associated with an immunosuppressive T cell phenotype. However, we observed a decrease in the infiltration of Tregs in high risk colitis, which seemed to peak again in dysplastic lesions.

Moreover, we assessed how β -1,6-GlcNAc branched *N*-glycans impact the interaction between T cells and cancer cells. Cocultures of T cells treated with kifunensine (KF), a glycosylation inhibitor, showed that the blockade of complex branched *N*-glycans synthesis promotes a T cell phenotype with increased cancer cell targeting efficiency, while *N*-acetylglucosamine (GlcNAc) treated T cells (which display higher levels of surface β -1,6-GlcNAc branched *N*-glycans) are considerably less efficiency in cancer surveillance. These results also highlight the importance of glycosylation, in particular of β -1,6-GlcNAc branched *N*-glycans, in the modulation of immune surveillance and response against CRC.

Taken together, our results suggest that overexpression of β -1,6-GlcNAc branched *N*-glycans in T cells is a mechanism responsible for the immunosuppression that leads to the promotion and development of serrated and Lynch syndrome CRC and CAC. This preliminary new data pinpoints a promising use of T cell glycoreprogramming as a potential directed therapy option for patients at risk of developing CRC or that have already been diagnosed with advanced stages.

Keywords: Cancer; Colorectal cancer; Inflammatory bowel disease; Colitis-associated cancer; *N*-glycosylation; β -1,6-GlcNAc branched *N*-glycans.

Resumo alongado da dissertação

O cancro colorretal é a designação de um grupo de neoplasias que surgem no cólon e reto e que são um grave problema de saúde pública. Visto que a sua incidência e mortalidade têm aumentando ao longo das últimas décadas, é crucial o estudo e compreensão desta patologia. A sobrevivência dos pacientes depende significativamente da deteção precoce. Embora melhorias nos métodos de rastreio e a deteção da doença em estadios iniciais tenham contribuído para uma diminuição da mortalidade, vários fatores, tais como a adesão a programas de rastreio, acesso a infraestruturas e recurso a cuidados de saúde de qualidade, influenciam no sucesso do diagnóstico precoce e no prognóstico dos pacientes. As estatísticas indicam que apenas cerca de 40% dos pacientes são diagnosticados com a doença em fases iniciais. O diagnóstico da doença em estadio avançado (cancro colorretal metastático) reduz significativamente a sobrevivência dos pacientes para cerca de 12%. Esta redução na sobrevivência está também associada à falta de eficácia terapêutica. Atualmente, cirurgias, quimioterapia e radioterapia são os procedimentos padrão para o tratamento destes pacientes. Apesar de novos tratamentos, como as imunoterapias, estarem continuamente a ser desenvolvidos, não são elegíveis para um número elevado de pacientes e apresentam reduzida eficiência no controlo da doença. Desta forma, existe uma enorme necessidade de identificar os mecanismos subjacentes ao desenvolvimento desta patologia de forma a descobrir novas terapias e melhorar as existentes para que seja possível reduzir o impacto do cancro colorretal e melhorar a qualidade de vida dos pacientes.

O cancro colorretal pode ocorrer esporadicamente ou em aglomerados familiares devido à transmissão de alelos mutantes aos descendentes (cancro colorretal hereditário). Além disso, esta é uma patologia altamente heterogénea, definida por grupos de alterações genéticas e epigenéticas distintas, designadas por vias de carcinogénese colorretal. Estas vias resultam em subtipos distintos de cancro colorretal que são molecular e fenotipicamente distintos uns dos outros. Particularmente importante no âmbito deste trabalho, o fenótipo do metilador das ilhas CpG (CIMP) e a instabilidade dos microssatélites (MSI) são as duas principais vias implicadas no cancro colorretal serrado. Este subtipo é caracterizado por lesões com uma morfologia "serrada" e por lesões progressivas que não seguem a sequência convencional adenoma-carcinoma. Em vez disso, o cancro colorretal serrado é explicado pela progressão de adenomas serrados não-displásticos para adenomas serrados displásticos por metilação ou perda do gene *MLH1*, sendo que estas últimas lesões progridem rapidamente para adenocarcinomas serrados. Entretanto, a MSI é responsável pela síndrome de Lynch, uma forma hereditária de cancro colorretal

que envolve a perda de genes codificantes de proteínas de reparação do genoma. O cancro colorretal serrado e associado à síndrome de Lynch são caracterizados por uma intensa infiltração imune. Além disso, a detecção de lesões precursoras nestes subtipos é bastante difícil, o que compromete a sobrevivência dos pacientes a longo prazo. Isto realça a necessidade urgente de estudar lesões pré-malignas com o objetivo de desenvolver novos alvos para prevenir a progressão da doença.

A doença inflamatória intestinal (DII) é uma doença caracterizada pela inflamação persistente do trato gastrointestinal, afetando severamente a qualidade de vida dos pacientes. A DII pode ser subdividida na colite ulcerosa (CU) e na doença de Crohn (DC). Atualmente, não existe nenhuma cura para a DII, e as terapias existentes carecem de eficácia na remissão completa da doença. Além disso, os doentes com DII têm um risco acrescido de desenvolver malignidades no intestino, designadas como cancro colorretal associado a colite (CAC), que está associado à extensão e duração da doença. Uma vez que o fardo da DII está a aumentar e os atuais métodos de rastreio não são totalmente eficazes na detecção de lesões pré-malignas em estádios iniciais, é possível que se observe uma maior incidência de CAC. Por esta razão, é fundamental melhorar o nosso conhecimento acerca desta cascata para desvendar soluções para este problema emergente.

O sistema imune é um importante mecanismo de defesa programado para proteger o hospedeiro contra qualquer ameaça. A perturbação da homeostase do sistema imunitário pode levar a várias doenças, nomeadamente as doenças autoimunes e cancro. Na verdade, as células cancerígenas empregam uma miríade de mecanismos que permitem a evasão da vigilância imunológica.

A *N*-glicosilação é uma modificação pós-traducional muito frequente, ocorrendo em cerca de 90% de todas as glicoproteínas. Os *N*-glicanos surgem da via secretora, que forma três tipos distintos de estruturas: *N*-glicanos ricos em manose, híbridos e complexos. Os *N*-glicanos são particularmente importantes no desenvolvimento e na atividade das células T. Curiosamente, a *N*-glicosilação desempenha um papel importante na mediação da patogénese da DII. As células T destes pacientes apresentam níveis reduzidos de *N*-glicanos β -1,6-GlcNAc ramificados no recetor da célula T, o que está associado a uma redução do limiar mínimo de ativação das células T e uma resposta inflamatória exacerbada. Ademais, o nosso grupo também já demonstrou que as células de cancro colorretal apresentam um glicoma característico, exibindo de forma aberrante *N*-glicanos β -1,6-GlcNAc ramificados. Demonstramos igualmente que a reprogramação destas células de forma a expressarem *N*-glicanos ricos em manose potencia uma resposta imune contra o tumor.

Neste trabalho, pretendemos descobrir qual o impacto dos *N*-glicanos β -1,6-GlcNAc ramificados na atividade das células T ao longo das diferentes cascatas de cancro colorretal: serreado, síndrome de Lynch e CAC. Usando coortes de pacientes bem caracterizadas (incluindo parafinas e biópsias frescas), encontramos evidências que sugerem que os *N*-glicanos β -1,6-GlcNAc ramificados são sobre expressos nas células T em lesões pré-malignas da via serreada e de síndrome de Lynch, o que por sua vez pode ser um fator responsável pela progressão da doença. Os nossos resultados preliminares de CAC também sugerem que estes glicanos complexos são sobre expressos nos estadios de colite de risco intermedio e elevado, o que aparenta estar associado com o risco de progressão de colite para displasia. Além disso, este perfil de glicosilação parece estar correlacionado com uma maior infiltração de células T regulatórias (Tregs), e sugere uma associação entre alterações no glicoma das células T e a criação de um ambiente imunossupressor propício para a proliferação de células cancerígenas. No entanto, observamos uma diminuição na infiltração de Tregs no estadio de colite de risco elevado, que apresentou simultaneamente os resultados mais elevados de expressão de *N*-glicanos β -1,6-GlcNAc ramificados.

Além disso, avaliamos o impacto dos *N*-glicanos β -1,6-GlcNAc ramificados na interação entre as células T e células cancerígenas. As coculturas com células T tratadas com kifunensina (KF), um inibidor da glicosilação, mostraram que o bloqueio da síntese de *N*-glicanos ramificados promove o desenvolvimento de células T mais eficazes na destruição de células cancerígenas, enquanto as células T tratadas com *N*-acetilglucosamina (GlcNAc) (que apresentam níveis mais elevados de *N*-glicanos β -1,6-GlcNAc ramificados) são consideravelmente menos eficientes contra células cancerígenas. Estes resultados também destacam a importância da glicosilação, em particular dos *N*-glicanos β -1,6-GlcNAc ramificados, na modulação da imunovigilância e resposta contra células malignas.

Os nossos resultados sugerem que a sobre expressão de *N*-glicanos β -1,6-GlcNAc ramificados em células T é um mecanismo responsável pelo aumento da capacidade imunossupressora destas células, sendo um fator importante na progressão de lesões pré-malignas para cancro (tal como avaliado nas cascatas de cancro colorretal serreado, síndrome de Lynch e CAC). Estes dados preliminares sugerem demonstram também que a manipulação do glicoma das células T pode ser uma eficiente forma de terapia direcionada para pacientes com risco de desenvolver cancro colorretal ou já tenham a doença em estadios mais avançados.

Palavras-chave: Cancro; Cancro colorretal; Doença inflamatória intestinal; Cancro associado à colite; *N*-glicosilação; *N*-glicanos β -1,6-GlcNAc ramificados

Table of contents

Agradecimentos.....	III
Extended dissertation synopsis	V
Resumo alongado da dissertação	IX
Table of contents	XIII
Figure Index	XV
Table Index.....	XXII
Abbreviation Index.....	XXIII
1. Introduction.....	1
1.1. Cancer.....	1
1.1.1. Colorectal cancer.....	2
1.1.1.2. Inflammatory bowel disease and colitis-associated colorectal cancer	7
1.2. The immune system	8
1.2.1. Tumor cells evade immune surveillance	10
1.3. Glycosylation	12
1.3.1. N-linked glycosylation	13
1.3.2. The glycosylation impact on T cell activity and disease.....	15
2. General aim	19
3. Materials and methods	21
3.1. Characterization of patient cohorts from CAC and CRC pathways	21
3.2. Lectin histochemistry	22
3.3. Immunohistochemistry.....	23
3.4. RNA extraction from FFPE samples	24
3.5. Cdna synthesis and real-time quantitative polymerase chain reaction (RT-Qpcr) ..	25
3.6. Fresh colonic biopsy and blood processing and FACS analysis	26
3.7. Cell lines and cell culture conditions	30
3.8. Splenocytes supplementation with kifunensine and N-acetylglucosamine	30
3.9. Splenocytes cocultures with MC38 cells – cytotoxic assay	32

3.10. Statistical analysis	33
4. Results	35
4.1. Characterization of the <i>in situ</i> immuno glycoprofile along distinct colorectal cancer cascades.....	35
4.1.1. Serrated and Lynch syndrome CRC premalignant and cancer lesions are characterized by an increased <i>lamina propria</i> immune infiltrate that overexpresses β -1,6- <i>N</i> -glycans	35
4.1.2. β -1,6 GlcNAc branched <i>N</i> -glycans expression appear to increase along the risk of progressing to colitis-associated colorectal cancer which is accompanied with an increase of infiltrate regulatory and cytotoxic T cells	45
4.1.3. Mice subjected to AOM/DSS induced colitis-associated cancer exhibit a similar glycoprofile observed in human patients.....	49
4.2. <i>Ex vivo</i> analysis of T lymphocyte glycoprofile in fresh biopsies from serrated and CAC pathways colorectal cancer patients.....	53
4.2.1. Expression of complex <i>N</i> -glycans is altered in T cells along the serrated cascade – glycosylation as a modulator of the T cell-mediated immune response favoring tumorigenesis.....	53
4.2.2. Increased expression of branched <i>N</i> -glycans and decreased sialylation contribute for an immuno suppressive environment in the progression to dysplasia	56
..... Error! Bookmark not defined.	
4.3. T lymphocyte glycoreprogramming: metabolic supplementation inducing overexpression of β -1,6- <i>N</i> -glycans inhibits effective immune response against cancer cells.....	60
5. Discussion	65
6. Conclusion and future perspectives	71
7. References	75

Figure Index

Figure 1 | Pie chart distribution of cancer cases worldwide in 2020, both sexes included. As evidenced, colorectal cancer (CRC) is the third most diagnosed cancer, after breast and lung cancer. From ¹⁵. _____ 3

Figure 2 | Estimated age-standardized incidence rates of colorectal cancer (CRC) per 100 000 inhabitants, both genders considered. Although developed regions (such as America, Australia, and Europe) currently present the highest number of cases, the trend over the years is for a stabilization or decrease of incidence. In the meantime, incidence in developing countries has been increasing. From ¹⁵. _____ 3

Figure 3 | Colorectal cancer (CRC) is the second most deadly malignancy, after lung cancer. Both genders were considered. These statistics demonstrate that CRC is a global public health issue that needs to be urgently addressed. From ¹⁵. _____ 4

Figure 4 | Colorectal cancer (CRC) molecular pathways. Lesions of chromosome instability (CIN) and microsatellite instability (MSI) pathways usually follow the adenoma-carcinoma sequence, which explains the majority of CRC cases. The MSI pathway is also characterized by a group of molecular changes associated with Lynch syndrome. Loss of the adenomatous polyposis coli (*APC*) and mismatch repair (*MMR*) genes lead to formation of small/early adenomas in the colon lining, growing to large/late adenomas. Mutations in apoptosis control genes and tumor growth factor β (*TGF β*) further lead to the progression to cancer. Meanwhile, the serrated pathway is characterized by phenotypically distinguishable lesions, grouped in the CpG island methylator phenotype (CIMP). This pathway involves hypermethylation of the DNA and loss of v-raf murine sarcoma viral oncogene homolog B1 (*BRAF*) which lead to sessile serrated adenomas (SSA) without dysplasia. Transition to SSA with dysplasia involve mutations of the MutL homolog 1 (*MLH1*) gene and can rapidly progress to serrated carcinoma by *p53* loss, further DNA methylation and other mutations. The last pathway is characterized by the presence of an inflammatory environment in the colon (such as those observed in IBD patients). This inflammatory pathway also presents an inversed order of mutations in *APC* and *p53* genes comparatively with CIN pathway. Based on ¹¹. _____ 6

Figure 5 | Cancer immunoediting phases. Tumor cells can be eliminated by the immune system through activation of innate and adaptive response (elimination phase). In the equilibrium phase, tumor cells are kept under pressure by the immune system, cycling between dormancy and division. Once they accumulate the necessary mutations, they can easily escape the immune cells, leading to disease progression (escape phase).

Mechanisms employed by tumor cells to suppress the immune system can be exploited to develop novel therapies. From ⁵⁶. _____12

Figure 6 | Different types of glycosylation that occurs in human cells. Glycans are sugar structures found ubiquitously in all life forms. Several distinct types of glycans exist, such as N-glycans, O-glycans, proteoglycans, glycosphingolipids, glycosylphosphatidylinositol (GPI)-linked proteins. From ⁶². _____13

Figure 7 | Simplified N-glycan biosynthesis pathway on the secretory pathway. Glycans are subject to the action of several enzymes, which add and remove sugar residues to the core structure. The pathway also relies on the action of important transport proteins. N-glycans can further be classified as oligomannose, hybrid or complex. Adapted from ⁶¹. _____15

Figure 8 | Glycans modulate several T cell mechanisms. The complex process of glycosylation is a key regulator of T cell homing to the thymus, development and selection, differentiation, proliferation, and homeostasis. Glycans also participate in immune tolerance, suppression, and autoreactivity observed in several diseases. From ⁷¹. _____16

Figure 9 | Blood components gradient separation with Lymphoprep™. Lymphoprep solution was pipetted into a falcon followed by the addition of peripheral blood without mixture, forming two layers. The sample was centrifuged at 2800 rpm for 20 minutes at 23°C. Afterwards, 4 distinctive layers could be distinguished based on the density of the components: an erythrocyte layer on the bottom, followed by the Lymphoprep layer, a layer containing the peripheral blood mononuclear cells (PBMCs) and the plasma layer. _____28

Figure 10 | General cytometry gating strategy for human biopsies and blood. The analysis started with the selection of cells based on the relation between size (forward scatter or FSC) and complexity/granularity (side scatter or SSC), as demonstrated by **A**. Afterwards, doublets cells were excluded by plotting event height (forward scatter height or FSC-H) by event area (forward scatter area or FSC-A) (**B**). Then live single cells negative for fixable viability dye (FVD) were selected (**C**). As the biopsies are composed of immune populations but also epithelial cells, cells expressing CD45 immune marker were selected (**D**). These populations were then divided on the several immune populations based on the expression of determined immune markers. As demonstrated in this example (**E**), helper T cells were selected by gating cells positive for CD4 immune marker. In this population, regulatory T cells were further selected by gating cells positive for FoxP3 immune marker (**G**) The glycoprofile was then determined by analyzing lectin-glycan interactions (**F** and **H**). _____30

Figure 11 | Effects of kifunensine (KF) and N-acetylglucosamine (GlcNAc) supplementation in T cell glycome. Kifunensine (KF) is a powerful inhibitor of the α -mannosidase I enzyme, blocking the N-glycosylation pathway to form more complex N-

glycans and promoting high-mannose structures. N-acetylglucosamine (GlcNAc) is a supplement which can power the hexosamine pathway and lead to the biosynthesis of uridine diphosphate N-acetylglucosamine (UDP-GlcNAc). UDP-GlcNAc is a substrate of the secretory pathway, leading to the assembly of complex branched N-glycans. _____ 32

Figure 12 | Evaluation of glycan and immune profile of serrated CRC pathway precursor lesions and carcinoma. Serrated pathway premalignant and carcinoma lesions were subjected to L-PHA and GNA lectin histochemistry for glycosylation profile evaluation, as well as immunohistochemistry for T cell co-receptor CD3, regulatory T cell transcription factor FoxP3 and cytotoxic T cell co-receptor CD8 to assess the immune infiltrate phenotypes. **(A)** Image panel representative of the lectin histochemistry - L-PHA and GNA (400x magnification; scale bar = 25 μ m). **(B)** Image panel representing immunohistochemistry - CD3, FoxP3 and CD8 (400x magnification; scale bar = 25 μ m). **(C)** Results from semi-quantification evaluation of L-PHA and GNA expression in the immune infiltrate of serrated lesions (performed by two independent evaluators) reveal glycome alterations throughout the pathway. **(D)** Results from semi-quantification of CD3, FoxP3 and CD8 markers in the immune infiltrate of serrated lesions (performed by two independent evaluators). The serrated cascade cohort was composed by sessile serrated adenomas without dysplasia (n=5), sessile serrated adenomas with dysplasia (n=5) and serrated CRC (n=3). Normal adjacent colonic tissue was used as control (n=2-7)_____ 38

Figure 13 | Evaluation of the glyco immune profile by Real time-quantitative polymerase chain reaction (RT-qPCR) of serrated cascade precursor lesions and carcinoma. FFPE samples representative of premalignant and carcinoma lesions of the serrated CRC pathway were subjected to RNA extraction, followed by conversion into cDNA and RT-qPCR amplification. Each sample analysis was performed in duplicates and gene expression was normalized for 18S housekeeping. Groups with $n \geq 3$ were tested for the presence of outlier values by the Robust regressing and Outlier removal (ROUT) method based on estimation of a maximum False Discovery Rate (FDR) of 1% (Q-value=1%). Statistical analysis was performed with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests: * $p \leq 0.05$; ** $p \leq 0.01$. **(A1)** *MGAT5* glycogene expression along the serrated cascade, normalized for *18S* and *CD3*, respectively **(A2)** *ST6GAL1* glycogene expression along the serrated cascade, normalized for *18S* and *CD3*, respectively. **(B)** Evaluation of immune related gene expression: *CD3* T lymphocyte common co-receptor, regulatory T lymphocyte transcription factor *FoxP3* and *CD8* cytotoxic T lymphocyte co-receptor, all normalized for *18S* housekeeping gene. Serrated cascade cohort composed by sessile serrated adenomas without dysplasia (n=20-22), sessile serrated adenomas with dysplasia (n=5-8) and serrated CRC (n=3). Normal colonic samples were included as controls (n=3-4)_____ 40

Figure 14 | Evaluation of glycan and immune profile of Lynch syndrome precursor lesions and CRC. Precursor and carcinoma lesions from Lynch syndrome patients were subjected to L-PHA and GNA lectin histochemistry for glycosylation profile evaluation, as well as immunohistochemistry for T cell co-receptor CD3, regulatory T cell transcription factor FoxP3 and cytotoxic T cell co-receptor CD8 to assess the immune infiltrate phenotypes. **(A)** Panel of images taken from representative slides (400x magnification; scale bar=25 μ m) demonstrating alterations in cell glycome - L-PHA and GNA. **(B)** Image panel representing final results from immunohistochemistry (images taken with 400x magnification; scale bar=25 μ m) - CD3, FoxP3 and CD8. **(C)** Results from semi-quantification evaluation of L-PHA and GNA expression in the immune infiltrate of serrated lesions (performed by two independent evaluators) point to alterations in the glycoprofile of immune cells **(C)** Results from semi-quantification of CD3, FoxP3 and CD8 immune markers on the infiltrate of Lynch syndrome lesions (performed by two independent evaluators). Lynch syndrome cascade cohort was composed of low grade dysplasia (LGD) lesions (n=6), high grade dysplasia (HGD) lesions (n=2) and Lynch CRC (n=3). Normal colonic samples were included as controls (n=2-7) _____ 42

Figure 15 | Evaluation of the glyco immune profile by Real time-quantitative polymerase chain reaction (RT-qPCR) of Lynch syndrome premalignant and carcinoma lesions. RNA was extracted from FFPE samples of precursor and Lynch carcinoma, converted into cDNA and analyzed by RT-qPCR . Each sample analysis was performed in duplicates and gene expression was normalized for 18S housekeeping. Groups with $n \geq 3$ were tested for the presence of outlier values by the Robust regressing and Outlier removal (ROUT) method based on estimation of a maximum False Discovery Rate (FDR) of 1% (Q-value=1%). Statistical analysis was performed with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests: * $p \leq 0.05$; ** $p \leq 0.01$. **(A1)** *MGAT5* glyco gene expression along the serrated cascade, normalized for 18S and CD3, respectively. **(A2)** *ST6GAL1* glyco gene expression along the serrated cascade, normalized for 18S and CD3, respectively. **(B)** Results of immune related gene expression: CD3 T lymphocyte common co-receptor, regulatory T lymphocyte transcription factor FoxP3 and CD8 cytotoxic T lymphocyte co-receptor, all normalized for 18S housekeeping gene. Lynch syndrome CRC cascade cohort was composed by low grade dysplasia (LGD) lesions (n=4-5), high grade dysplasia (HGD) lesions (n=2) and Lynch CRC (n=2-3). Normal colonic samples were included as controls (n=2-4). _____ 44

Figure 16 | Lectin histochemistry analysis of the CAC cascade immune infiltrate glycosignature. **(A)** Image panel representing the glycoprofile observed along the CAC cascade (400x magnification; scale bar = 25 μ m) **(B)** Stromal semi-quantification results from L-PHA and GNA lectin histochemistry performed on lesions from the CAC cascade:

healthy control (n=7), non-inflamed tissue (includes adjacent tissue and tissue with macroscopic remission) (n=26), low risk colitis (n=26), intermediate risk colitis (includes polyps and Mayo 2 and 3 inflammation samples) (n=22), high risk colitis (pancolitis) (n=7), low grade dysplasia (n=3) and carcinoma (n=1). Semi-quantification was performed by two independent evaluators. _____ 46

Figure 17 | Evaluation of the glyco immune profile by Real time-quantitative polymerase chain reaction (RT-qPCR) of CAC cascade lesions. Each sample analysis was performed in triplicates and gene expression was normalized for 18S housekeeping. Groups with $n \geq 3$ were tested for the presence of outlier values by the Robust regressing and Outlier removal (ROUT) method based on estimation of a maximum False Discovery Rate (FDR) of 1% (Q-value=1%). Statistical analysis was performed with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests. **(A)** *MGAT5*, *MAN2A1*, *FUT8* and *ST6GAL1* glycosyltransferases expression along CAC cascade. **(B)** Expression of immune genes coding *FoxP3* transcription factor and *CD8* cytotoxic T cell co-receptor. CAC cascade cohort composed by non-inflamed tissue (includes adjacent tissue and tissue with macroscopic remission) (n=3-4), low risk colitis (n=2), intermediate risk colitis (includes polyps and Mayo 2 and 3 inflammation samples) (n=3-17), high risk colitis (pancolitis) (n=3), low grade dysplasia (n=3) and carcinoma (n=1). Normal colonic samples were used as healthy controls (n=3-4). _____ 48

Figure 18 | Evaluation of glycoprofile by GNA lectin histochemistry performed on AOM/DSS treated WT and *Mgat5*^{-/-} mice. **(A)** Image panel representative of GNA signal on the immune infiltrate of a series of samples from wild type mice subjected to AOM/DSS (400x magnification; scale bar =25 μ m). **(B)** Image panel representative of GNA signal on the immune infiltrate of a series of samples from *Mgat5*^{-/-} mice subjected to AOM/DSS (400x magnification; scale bar =25 μ m). **(C)** Semi-quantification results of expression of high mannose N-glycans along CAC pathogenesis in Wild type mice. **(D)** Semi-quantification results of expression of high mannose N-glycans along CAC pathogenesis in *Mgat5*^{-/-} mice. Groups with $n \geq 3$ were analyzed for statistical significance with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests: * $p \leq 0.05$. AOM/DSS- treated WT mice cohort was composed of normal controls (n=6), acute inflammation (n=3), chronic inflammation (n=5), low grade dysplasia (LGD) (n=3), high grade dysplasia (HGD) (n=4) and colitis-associated colorectal cancer (CAC) (n=3). AOM/DSS-treated *Mgat5*^{-/-} mice cohort was composed of normal controls (n=5) acute inflammation (n=3), chronic inflammation (n=3), LGD (n=3), high grade dysplasia (HGD) (n=6) and CAC (n=3) _____ 51

Figure 19 | Evaluation of the glycoimmune profile in premalignant serrated precursor lesions and CRC by FACS. **(A)** Percentage of CD45⁺ CD4⁺ T cells (gated on live single

cells). MFI results from L-PHA and GNA lectin staining and L-PHA/GNA ratio. **(B)** Percentage of CD45⁺ CD4⁺ FoxP3⁺ CD25⁺ regulatory T cells (gated on live single cells). MFI results from L-PHA and GNA lectin staining and L-PHA/GNA ratio. **(C)** Percentage of CD45⁺ CD8⁺ T cells (gated on live single cells). MFI results from GNA lectin staining. Groups with $n \geq 3$ were tested for the presence of outlier values by the Robust regression and Outlier removal (ROUT) method based on estimation of a maximum False Discovery Rate (FDR) of 1% (Q-value=1%). Statistical analysis was performed with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests: $*p \leq 0.05$. The cohort of serrated lesions used for FACS analysis included sessile serrated adenomas without dysplasia (n=8), sessile serrated adenomas with dysplasia (n=2) and CRC samples (n=2). Normal colonic samples were used as controls (n=6) _____ 55

Figure 20 | CD4⁺ T cells glycoprofile in CAC fresh biopsies. **(A)** Percentage of infiltrating CD45⁺ CD4⁺ T cells in colon lesions. L-PHA, GNA and SNA lectin staining MFI of CD4⁺ T cells. **(B)** Percentage of circulating CD45⁺ CD4⁺ T cells. L-PHA, GNA and SNA lectin staining MFI of blood CD4⁺ T cells. No statistical analysis was performed due to the low sample size. CAC cascade stages include remission/non-inflamed adjacent tissue (n=3 for biopsies and n=1 for blood) and dysplasia lesions (n=1 for biopsies and blood) _____ 58

Figure 21 | CD8⁺ T cells glycoprofile in CAC fresh biopsies. **(A)** Percentage of infiltrating CD45⁺ CD8⁺ T cells in colon lesions. L-PHA, GNA and SNA lectin staining MFI of CD8⁺ T cells. **(B)** Percentage of circulating CD45⁺ CD8⁺ T cells. L-PHA, GNA and SNA lectin staining MFI of blood CD8⁺ cytotoxic T cells. No statistical analysis was performed due to the low sample size. CAC cascade stages include remission/non-inflamed adjacent tissue (n=3 for biopsies and n=1 for blood) and dysplasia lesions (n=1 for biopsies and blood) 59

Figure 22 | Metabolic supplementation and T cell reprogramming optimization experiment. **(A)** L-PHA staining (FITC MFI), **(B)** GNA staining (FITC MFI) and **(C)** ConA staining (biotinylated lectin detected by streptavidin- PE MFI) of T cells cultured with 20 μ M of Kifunensine for 24, 48 and 48 hours + 24 hours of recovery compared with untreated splenocytes. **(D)** L-PHA staining (FITC MFI) of T cells cultured in medium supplemented with 100 mM of GlcNAc for 24, 48 and 48 hours + 24 hours of recovery compared with untreated control. No statistical analysis was performed due to the low sample size ($n \leq 3$). _____ 61

Figure 23 | MC38 murine adenocarcinoma cells viability after coculture with splenocytes untreated or treated with kifunensine. Cocultured cells were stained for FACS analysis. Cancer cells were distinguished from immune cells by CD45 staining and evaluated for cell death by annexin-V/PI method in order to assess the efficacy of splenocyte immune response after glycoreprogramming. No statistical analysis was performed. **(A)** Graphic plot of apoptosis assay percentages. **(B)** Mean percentages of each

population distinguished by annexin-V/PI. This experiment was conducted with MC38 control cells (n=3), splenocytes + MC38 cocultures (n=2) and KF-treated splenocytes + MC38 cocultures (n=3). _____ 62

Figure 24 | MC38 murine adenocarcinoma cells viability after coculture with splenocytes untreated or treated with N-acetylglucosamine. Cocultured cells were stained for FACS analysis. Cancer cells were distinguished from immune cells by CD45 staining and evaluated for cell death by annexin-V/PI method in order to assess the efficacy of T cell response after glycoreprogramming. **(A)** Graphic plot of apoptosis assay percentages. Statistical analysis was performed by an ordinary two-way ANOVA with multiple comparisons: *** $p \leq 0.001$; **** $p \leq 0.0001$. **(B)** Mean percentages of each population distinguished by annexin-V/PI. This experiment was conducted with MC38 control cells (n=3), splenocytes + MC38 cocultures (n=3) and GlcNAc-treated splenocytes + MC38 cocultures (n=3). _____ 63

Figure 25 | T cell glycoprofile in CAC cascade (progression from inflammation to premalignancy and finally to carcinoma) and in serrated and Lynch syndrome cascades (progression from premalignancy to carcinoma). **(A)** T cell N-glycan surface expression along the cascades. **(B)** Glycogene and immune gene expression changes relatively to homeostatic T cell profile. The CAC cascade preliminary results suggest an increase in expression of β -1,6-linked branched N-glycans on T cell surface along the inflammatory pathway according to colitis risk. Moreover, this seems to be associated with an increase infiltration of immunosuppressive cells. However, we observed a trend decrease in infiltration of these Tregs in the stage which branched N-glycans seem to peak. Serrated and Lynch syndrome cascades are characterized by over expression of these branched structures, which are associated with decreased cancer cell targeting. These glycome changes occur early, as we demonstrated. Finally, we demonstrated that glycoreprogramming T cells has an immense potential as a directed therapy for CRC resolution. _____ 73

Table Index

Table 1 | Characterization of lectins used in the present work. Lectins are glycan-binding proteins (GBPs) that recognize specific glycan domains. Glycans structures adapted from chapter 48 of the Essentials of Glycobiology ⁸⁰. _____23

Table 2 | Amplification probes used for gene expression quantification by RT-qPCR. All probes used are from TaqMan TM, except the 18S probe which is manufactured by Integrated DNA Technologies (IDT). _____26

Table 3 | Antibodies and lectins used for FACS analysis, with respective fluorochrome, dilution and manufacturer. _____28

Abbreviation Index

A

Adenomatous polyposis coli (APC)
allophycocyanin (APC)
Antigen-presenting cells (APCs)
Avidin-biotin complex (ABC)

B

B cell receptor (BCR)
Bovine serum albumin (BSA)

C

Cancro colorretal associado à colite (CAC)
Centro Hospitalar Universitário do Porto (CHUP)
Chemokine ligand 20 (CCL20)
Chromosomal instability (CIN)
Cluster of differentiation 4 (CD4⁺)
Cluster of differentiation 8 (CD8⁺)
Colite ulcerosa (CU)
Colitis-associated colorectal cancer (CAC)
Colorectal cancer (CRC)
Concanavalin A lectin from *Canavalia ensiformis* (ConA)
Coronavirus disease (COVID-19)
Cpg island methylator phenotype (CIMP)
Crohn's disease (CD)
Cyanine 5.5 (Cy5.5)
Cyanine 7 (Cy7)
Cytotoxic T lymphocytes (CTLs)

D

Dendritic cells (DC)
Deoxyribonucleotide triphosphate (DNTPs)
Dextran sodium sulfate (AOM/DSS)
Diaminobenzidine (DAB)
Doença de Crohn (DC)
Doença inflamatória intestinal (DII)
Dolichol-phosphate (DoI-P)

E

Endoplasmic reticulum (ER)
Extracellular matrix (ECM)

F

Familial adenomatous polyposis (FAP)
Fetal bovine serum (FBS)
Fixable viability dye (FVD)
Fluorescein isothiocyanate (FITC)
Fluorescence-activated cell sorting (FACS)
Forkhead Box P3 (*FoxP3*)
Formalin-fixed paraffin-embedded (FFPE)
Fucosyltransferase 8 (*FUT8*)

G

Galanthus nivalis (GNA)
GATA-binding protein 3 (GATA-3)
Glucose (Glc)
Glycan-binding proteins (GBPs)
Glycosylphosphatidylinositol (GPI)
Glycosyltransferases (GTs)
Golgi apparatus (GA)

Granulocyte-macrophage colony-stimulating factor (GM-CSF)

H

Hanks' Balanced Salt Solution (HBSS)
Hereditary nonpolyposis colorectal cancer (HNPCC)
High grade dysplasia (HGD)
Hyperplastic polyps (HP)

I

Immunohistochemistry (IHC)
Inflammatory bowel disease (IBD)
Innate lymphoid cells (ILCs)
Integrated DNA Technologies (IDT)
Interferon- γ (IFN- γ)
Interleukin-17 (IL-17),
Interleukin-17A (IL-17A)
Interleukin-17F (IL-17F)
Interleukin-2 (IL-2)
Interleukin-22 (IL-22)
Interleukin-23 (IL-23)
Interleukin-4 (IL-4)
Interleukin-6 (IL-6),

K

Kifunensina (KF)
Kifunensine (KF)

L

Low grade dysplasia (LGD)

M

Major histocompatibility complex I (MHC I)
Major histocompatibility complex II (MHC II)

Mannose (Man)
Median fluorescence intensity (MFI)
Messenger RNA (mRNA)
Microsatellite instability (MSI)
Mismatch repair (MMR)
Mutl homolog 1 (MLH1)
Myeloid-derived suppressor cells (MDSC)

N

N-acetylglucosamina (GlcNAc)
N-acetylglucosamine (GlcNAc)
N-Acetylglucosamine-phosphate (GlcNAc-P)
N-acetylglucosaminyltransferase I (GnT-I)
N-acetylglucosaminyltransferase IV (GnT-IV)
N-acetylglucosaminyltransferase IV (GnT-V)
Natural killer cells (NK)
Natural killer T cells (NKT)
Nod-like receptors (NLRs)

O

Oligosaccharyltransferase (OST)

P

Pacific Blue (PB)
Pathogen-associated molecular patterns (PAMPs)
PBS 1x + 0,01% Tween (PBST)
Penicillin/Streptomycin (Pen/Strep)
Peridinin chlorophyll protein (PerCP)
Peripheral blood mononuclear cells (PBMCs)

Phaseolus vulgaris Leucoagglutinin (L-PHA)

Phycoerythrin (PE)

Phosphate buffered solution (PBS)

Poly-*N*-acetyllactosamine (poly-LacNAc)

Dulbecco's modified Eagle's minimal essential medium (DMEM)

Programmed death ligand 1 (PD-L1)

Propidium iodide (PI)

P-selectin glycoprotein ligand-1 (PSGL-1)

R

RAR-related orphan receptor alpha (ROR α)

RAR-related orphan receptor gamma (ROR γ t)

Reactive oxygen species (ROS)

Real time- quantitative polymerase chain reaction (RT-qPCR)

Regulatory T cells (Tregs)

Relative quantification (RQ)

Room temperature (RT)

Roswell Park Memorial Institute (RPMI)

S

Sambucus nigra agglutinin (SNA)

Sessile serrated lesions (SSL)

Signal transducer and activator of transcription 6 (STAT-6)

ST6 β -galactoside α -2,6-sialyltransferase 1 (*ST6GAL1*)

T

T cell receptor (TCR)

T helper cells (Th)

T-box transcription factor *TBX21* (T-bet)

Thymus-seeding progenitors (TSPs)

Toll-like receptors (TLRs)

Traditional serrated adenomas (TSA)

Transforming growth factor- β (TGF- β)

Tumor necrosis factor- α (TNF- α)

Tumor necrosis factor- β (TNF- β)

U

Ulcerative colitis (UC)

Uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc)

V

V-raf murine sarcoma viral oncogene homolog B1 (BRAF)

W

Wild type (WT)

α

α -1,6-mannosylglycoprotein 6- β -*N*-Acetylglucosaminyltransferase A (*MGAT5*)

α -mannosidase II (*MAN2A1*)

$\alpha\beta$ T cell receptor ($\alpha\beta$ TCR)

1. Introduction

1.1. Cancer

Cancer is a leading public health issue. The rise in cancer cases is partially attributed to the increase in life expectancy that has been observed along the past decades. Furthermore, increased exposure to risk factors also explains why the rise in cancer cases positively correlates with socioeconomic development ^{1,2}. Cancer is a leading cause of death worldwide. In parallel to estimated incidence rates for 2040, the mortality rate is also expected to increase ¹. Nevertheless, it is important to note that the future of cancer incidence and mortality is highly uncertain due to the ongoing impact of the coronavirus disease (COVID-19) pandemic ^{1,3}. The sudden emergence of this novel public health issue had implications in cancer patient diagnosis and care (whether due to the need to divert health professionals to the treatment of COVID-19 patients or in order to decrease the exposure of patients to the virus in hospital environments), which revealed the fragility in health systems around the world, with a consequent decrease in the effectiveness of early patient diagnosis and treatment ³. For all these reasons, there is a great need to improve patient diagnosis and develop advanced therapies to effectively reduce the burden cancer patients present to national health services and to reduce the burden of disease in the general population.

Tumor development can be attributed to somatic mutations occurring in a cell type, which awards it with a selective advantage over others. Mutations in specific *loci* of oncogenes or suppressor genes can lead to either overexpression/hyperactivity or downregulation of specific proteins that are essential in the protection of genomic and cellular integrity, respectively ⁴. Tumor progression involves multistep events, namely different mutations that initiate tumorigenesis and others that accumulate and gradually confer an improved selective advantage for tumor cells to continue to grow and spread ⁵.

Cancer is a complex pathology. In recent decades of research, several characteristics have been proposed as key factors in the initiation and progression of carcinogenesis. These characteristics have been denominated as the hallmarks of cancer: traits that cells acquire that allow them to become increasingly adapted to their environment, while also promoting growth, proliferation and dissemination ⁶. These hallmarks include mechanisms used to evade several checkpoints responsible for coordinating cell function and maintaining homeostasis. Cancer cells are characterized by: dysregulated proliferative signaling (cells become independent of the environment they are inserted and proliferate without any control); ability to evade growth suppressor signaling; resistance to programmed cell death induction; activation of telomerase enzyme and immortalization;

induction of neoformation of blood vessels for nutrient and oxygen supply (angiogenesis); activation of invasion and metastasis mechanisms; metabolism reprogramming; immune system evasion; insertion in a tumor promoting pro-inflammatory microenvironment ; increased mutation rate and genomic instability ⁶.

Furthermore, the tumor microenvironment includes the transformed cells, nutrients, signaling factors, blood vessels, the extracellular matrix (ECM) and also non-transformed normal immune and stromal cells that associated to form an intricate framework that sustains malignant progression ^{7,8}.

Colorectal cancer (CRC) is a group of malignancies arising in the colon and rectum with high relevance in terms of incidence and mortality ⁹⁻¹¹. Moreover, CRC is a highly heterogeneous disease characterized by distinct pathways ¹². This pathology will be the main focus of the present thesis and further discussion regarding its characteristics and impact will be addressed in the next section.

1.1.1. Colorectal cancer

The burden of CRC is worrisome and future estimates are uncomfortable. Several factors are linked with an increased risk of developing CRC, including tobacco and alcohol abuse, diet, physical activity, age, race, family history, and presence of comorbidities ¹³. Prevalence of some of these risk factors are associated with the rapid economic development in the recent decades, correlating directly with the increase in CRC incidence and mortality.

Epidemiologically, statistics point CRC as the third most incident cancer worldwide (**Figure 1**) ⁹. When analyzed by gender, CRC is the third and second most diagnosed cancer in men and women, respectively ^{14,15}. Although incidence varies substantially worldwide, a significant increase in cases has been observed in low- to medium-income (which can be attributed to the introduction of western lifestyle habits associated with socioeconomic development). Meanwhile, high-income countries present the highest number of diagnosed CRC cases (**Figure 2**) ^{9,10,13,16}.

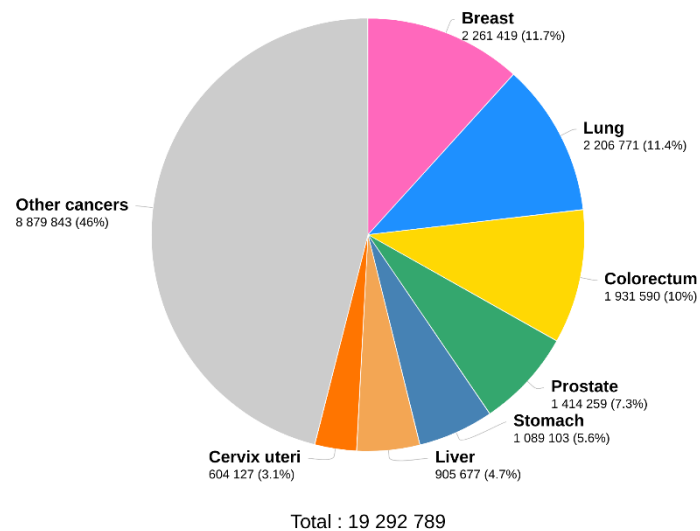


Figure 1 | Pie chart distribution of cancer cases worldwide in 2020, both sexes included. As evidenced, colorectal cancer (CRC) is the third most diagnosed cancer, after breast and lung cancer. From ¹⁵.

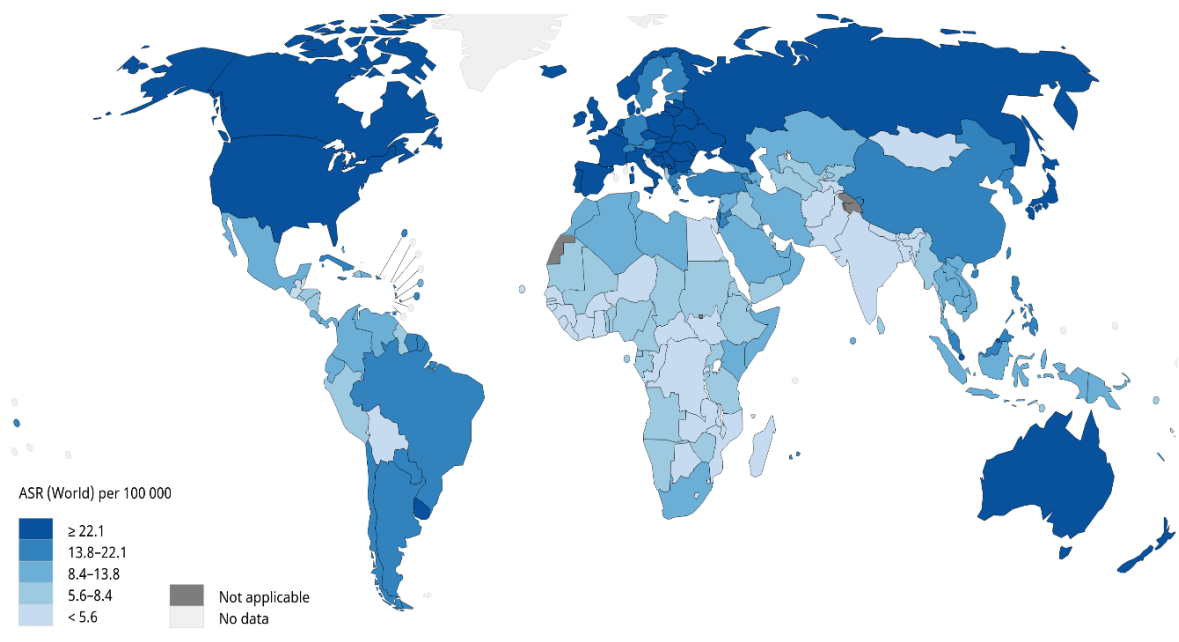


Figure 2 | Estimated age-standardized incidence rates of colorectal cancer (CRC) per 100 000 inhabitants, both genders considered. Although developed regions (such as America, Australia, and Europe) currently present the highest number of cases, the trend over the years is for a stabilization or decrease of incidence. In the meantime, incidence in developing countries has been increasing. From ¹⁵.

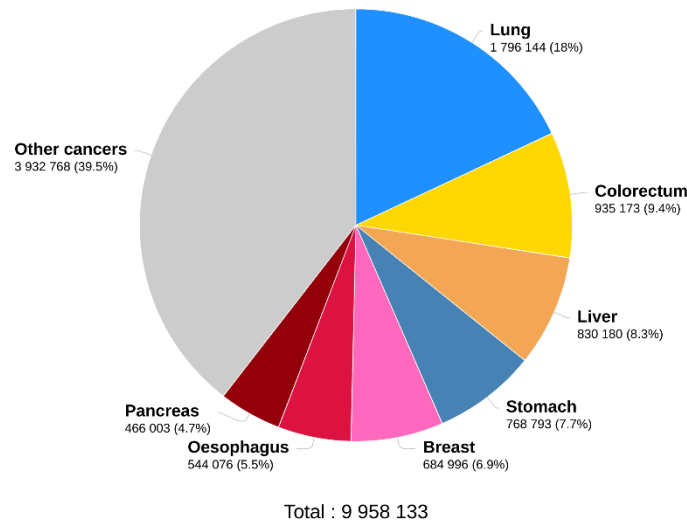


Figure 3 | Colorectal cancer (CRC) is the second most deadly malignancy, after lung cancer. Both genders were considered. These statistics demonstrate that CRC is a global public health issue that needs to be urgently addressed. From ¹⁵.

Regarding the mortality, CRC is the fourth most common cancer-related cause of death in men and third in women, respectively ¹⁰, although more recent statistics point to CRC as the second leading cancer-related mortality cause (**Figure 3**) ¹⁷. In general, developed countries have managed mortality rates through better designed screening programs, while developing countries, which lack economic power and necessary infrastructure, have more difficulty to effectively diagnosing the disease in earlier stages and successfully treating patients ^{10,11,16}. Combining these facts with the increase in CRC cases in these countries, it is not surprising that mortality rates are increasing at alarming rates. Particularly, around 1 quarter of patients with CRC are diagnosed with advanced metastatic disease and 5-year survival remains around 64% (which drops to 12% in cases of metastatic CRC) ^{14,18}. In addition, existing therapies are not always effective ¹⁴. For this reason there is a great need for improvement in screening methods and patient adherence, as well as detect novel biomarkers and therapeutic targets so that CRC cases can be detected in its initial stages¹⁹. Currently, a major focus has been placed on the potential of immunotherapies as a therapeutic option for CRC ^{20,21}. However, studies conducted have only demonstrated successful results by using checkpoint inhibitor immunotherapies in the CRC subtype characterized by loss of mismatch repair (MMR) genes and high microsatellite instability (MSI), while other CRC subtypes have either been neglected or produced unsatisfactory results ²¹.

Most CRC cases are sporadic (around 60-65%), which means that they occur spontaneously without the transmission of mutations throughout generations ²².

Nevertheless, around 25% of cases occur in familial clusters and 5% are due to heritable mutations that can cause cancer syndromes, including familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC) or Lynch syndrome)¹¹.

1.1.1.1. Sporadic and hereditary colorectal cancer pathways

Molecularly, CRC is a highly heterogeneous disease that involves sequential key gene mutations and epigenetic alterations leading to distinct phenotypes¹². Chromosomal instability (CIN) and MSI pathways involve mutations that lead to progressively more malignant lesions that follow the traditional adenoma-carcinoma sequence (**Figure 4**)^{11,12,23}. The CIN pathway actually explains the majority of sporadic cases of CRC and the MSI pathway is responsible for nearly all Lynch syndrome-associated CRC cases (a hereditary form of CRC)²³. MSI usually initiates by loss of the *adenomatous polyposis coli* (APC) gene, although mutations in *v-raf murine sarcoma viral oncogene homolog B1* (BRAF) can also promote the carcinogenic pathway¹². These mutations occur due to inherited errors in MMR genes²³ (**Figure 4**). An alternative and less studied pathway leading to CRC is named the serrated pathway (a subtype of sporadic cancer which doesn't follow the adenoma-carcinoma sequence) and involves BRAF mutations, MSI instability, oncogene activation, and hypermethylation of CpG islands in tumor suppressor gene promoter regions. This pathway is classified as a CpG island methylator phenotype (CIMP) subtype^{11,12,23,24}. In this pathway, precursor lesions can be classified as traditional serrated adenomas (TSA), hyperplastic polyps (HP) and sessile serrated lesions (SSL). SSL have a particular increased risk of progressing to malignancy, although its degree is unknown²⁴. SSL progress to cancer by intermediate dysplastic lesions (sessile serrated adenomas with dysplasia), a step associated with the silencing or loss of *MutL homolog 1* (MLH1) gene²⁵ (**Figure 4**). In fact, it is highly important to study and evaluate CRC precursor lesions in patients, since individuals harboring these precancerous lesions have an increased risk of developing CRC and their analysis can provide valuable insight about exact patient risk and their outcome²². A third pathway, whose occurrence is quite rare, is the inflammatory pathway of CRC, responsible for less than 2% of CRC cases¹¹ (**Figure 4**). This pathway will be discussed in the following chapter.

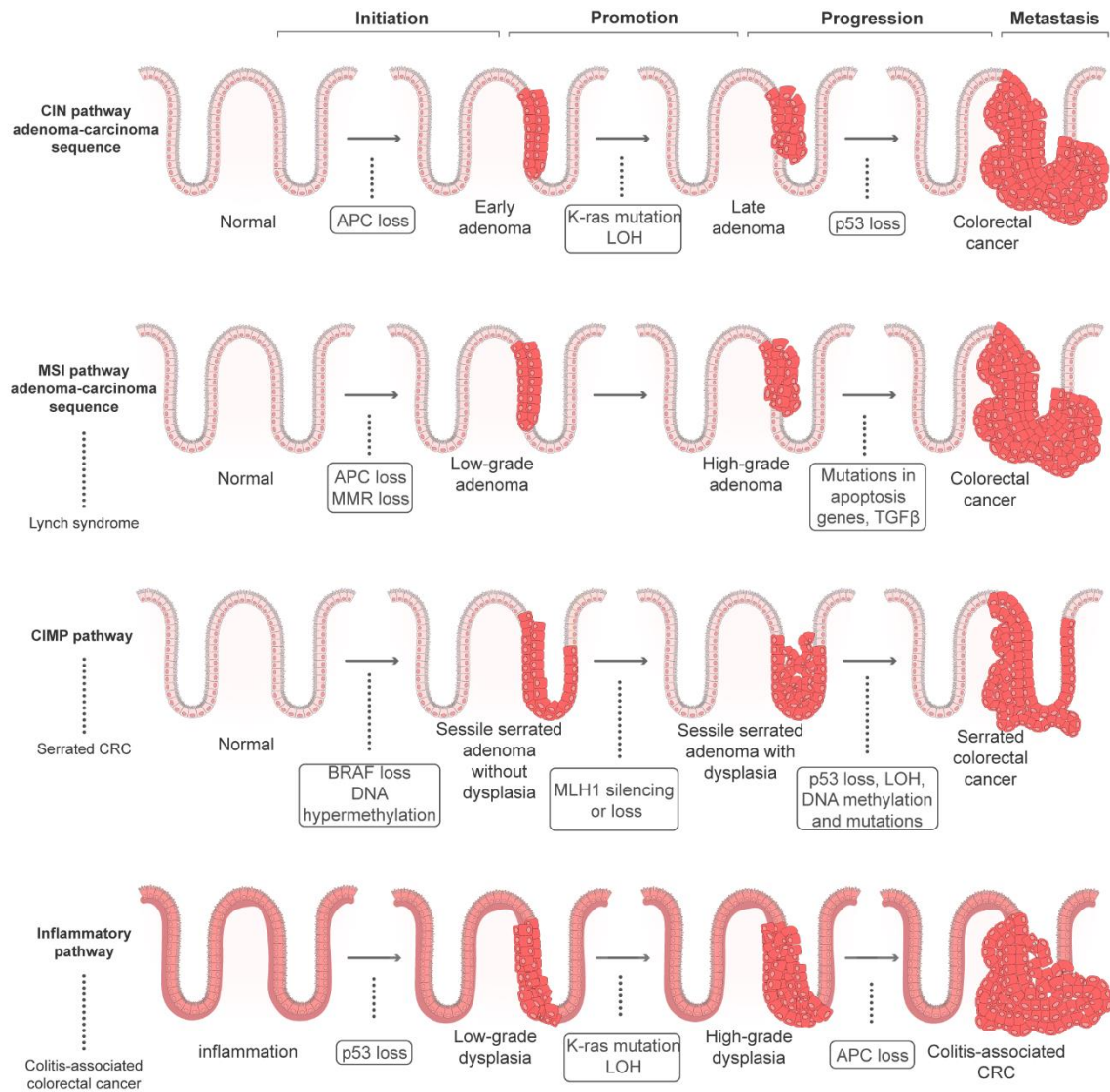


Figure 4 | Colorectal cancer (CRC) molecular pathways. Lesions of chromosome instability (CIN) and microsatellite instability (MSI) pathways usually follow the adenoma-carcinoma sequence, which explains the majority of CRC cases. The MSI pathway is also characterized by a group of molecular changes associated with Lynch syndrome. Loss of the *adenomatous polyposis coli* (*APC*) and mismatch repair (*MMR*) genes lead to formation of small/early adenomas in the colon lining, growing to large/late adenomas. Mutations in apoptosis control genes and *tumor growth factor β* (*TGFβ*) further lead to the progression to cancer. Meanwhile, the serrated pathway is characterized by phenotypically distinguishable lesions, grouped in the CpG island methylator phenotype (CIMP). This pathway involves hypermethylation of the DNA and loss of *v-rat murine sarcoma viral oncogene homolog B1* (*BRAF*) which lead to sessile serrated adenomas (SSA) without dysplasia. Transition to SSA with dysplasia involve mutations of the *MutL homolog 1* (*MLH1*) gene and can rapidly progress to serrated carcinoma by *p53* loss, further DNA methylation and other mutations. The last pathway is characterized by the presence of an inflammatory environment in the colon (such as those observed in IBD patients). This inflammatory pathway also presents an inversed order of mutations in *APC* and *p53* genes comparatively with CIN pathway. Based on ¹¹.

1.1.1.2. Inflammatory bowel disease and colitis-associated colorectal cancer

Inflammation is a normal biological mechanism induced by the immune system in response to infection and/or tissue damage. Uncontrolled or unresolved inflammation can have serious consequences, leading to diseases associated with chronic inflammation, such as inflammatory bowel disease (IBD) ²⁶, which is a lifelong disease whose incidence and prevalence have been increasing over the years, both in developed and developing countries ²⁷. IBD includes ulcerative colitis (UC), Crohn's disease (CD) and other pathologies with distinct characteristics. In fact, UC is characterized by a continuous pattern of inflammation along the mucosa and submucosa, while CD manifests a discontinuous pattern of inflammation which can affect any layer of the colon ²⁸. Along with genetic risk factors, several environmental factors, including tobacco abuse, diet, gut microbiome dysbiosis, and immune dysregulation, all contribute to an increased risk of intestinal inflammation ^{27,28}.

The persistent bowel inflammation presents immense pain to patients. Indeed, IBD is a highly debilitating disease that affects all aspects of the lives of patients (physical, psychological, and social) for which no cure is available. Furthermore, existing therapies are not fully effective and a significant percentage of patients fail to respond to treatment ^{29,30}. Moreover, the extensive immune response against commensal bacteria – non-pathogenic bacteria that reside in the gastrointestinal tract and form a symbiotic connection with the host - which leads to the exaggerated harmful inflammation also predisposes patients to develop colitis-associated colorectal cancer (CAC) ³¹. In fact, inflammation is associated with the production of reactive oxygen species (ROS), which are capable of inducing genomic and mitochondrial DNA damage and mutagenesis. Mitochondrial DNA damage is particularly important because it leads to impaired mitochondrial function and free radical generation, increasing cell instability, and increasing genomic DNA damage ³². In addition, epigenetic alterations, aberrations of the chromosome, inactivation of tumor suppressor genes, and activation of oncogenes all play an important role in the pathogenesis of CAC ^{31,32}.

CAC can occur at younger ages and is phenotypically distinct from other CRC pathways, since it does not follow the sequence from adenoma to carcinoma ^{29,33}, progressing instead from flat dysplastic lesions to carcinoma, which when combined with the extremely inflammatory environment and scarring of the epithelium results in lesions that are very difficult to detect ³⁴. With the increase in IBD cases, the world must be prepared for the possibility of CAC cases also increasing and must develop strategies to effectively detect

and treat these patients. Research is needed in order to discover new biomarkers and therapeutic targets that will solve this problem ³⁵.

1.2. The immune system

The immune system is a complex network of intrinsic pathways whose main objective is the defense of the host and the maintenance of homeostasis ^{36–38}. Nevertheless, the immune response also plays a key role in the several pathologies which endanger the host, for instance: autoimmunity, immunodeficiency and hypersensitivity/allergies ^{36–39}. In autoimmune diseases, the distinction between self and non self is compromised and leads to disruption of the immune system tolerance. Immune cells possess mechanisms controlling this tolerance and failure is responsible for an irregular immune response ³⁹. In addition, the immune system also plays an important role in fighting cancer cells ^{36,40,41}. Indeed, disruption of the immune system balance is a key hallmark of cancer ⁶.

Two main groups of immune response can be distinguished: innate and adaptive response. The innate response comprises several body barriers (including skin and mucous), temperature and pH, and several cell types - macrophages, dendritic cells (DCs), neutrophils, natural killer cells (NK), innate lymphoid cells (ILCs), natural killer T cells (NKT), $\gamma\delta$ -T cells, among others - whose working mechanism is more primitive ⁴². These cells recognize pathogenic microbes through specific membrane proteins, such as, toll-like receptors (TLRs) and Nod-like receptors (NLRs) which detect common pathogen-associated molecular patterns (PAMPs) ⁴³. This first line of defense is an earlier and nonspecific response, although it is usually enough to defend against most invading microbes ⁴⁴. The adaptive immune response is a more advanced evolutionary mechanism with high specificity and relies on the presentation of antigens to lymphocytes, a subgroup of leucocytes, namely T cells and B cells ⁴⁴.

T helper cells (Th) express the cluster of differentiation 4 (CD4⁺) and an $\alpha\beta$ T cell receptor ($\alpha\beta$ TCR). As their name suggests, these cells help polarize the immune response in accordance with the type of invading microbe. Th cells achieve this by releasing effector cytokines, which vary according to their subtype ⁴². Differential gene expression, mechanisms of gene silencing and exposure to select types of cytokines implicate differentiation into a specific Th cell subtype, which are further subdivided into Th1, Th2, Th17 and Tregs ⁴⁵. Th1 cells are characterized by the expression of the T-box transcription factor *TBX21* (T-bet) and secrete pro-inflammatory cytokines, including interferon- γ (IFN- γ), tumor necrosis factor- α (TNF- α) and tumor necrosis factor- β (TNF- β), interleukin-2 (IL-2) and granulocyte-macrophage colony-stimulating factor (GM-CSF). These cells are

essential in the response against intracellular pathogens ^{42,45–47}. Meanwhile, Th2 cell differentiation is controlled by GATA-binding protein 3 (GATA-3) transcription factor and is induced by the signal transducer and activator of transcription 6 (STAT-6) and interleukin-4 (IL-4). These cells are responsible for protection against extracellular pathogens and are secretors of IL-4, IL-5, IL-9, IL-10, and IL-13 ^{42,45–47}. A distinct subtype was identified years after the definition of Th1 and Th2 ⁴⁸. This subtype was classified as Th17. Th17 present low expression of T-bet and GATA-3 and high expression of RAR-related orphan receptor alpha (ROR α) and gamma (ROR γ t). These cells are major producers of interleukin-17A (IL-17A), interleukin-17F (IL-17F), interleukin-22 (IL-22) and chemokine ligand 20 (CCL20), playing a role in neutrophil recruitment and antimicrobial peptide production in response to a wide variety of pathogens ^{42,46–48}. Lastly, regulatory T cells (Tregs) express forkhead box P3 (Foxp3) transcription factor and are known for their suppressive function, regulating immune response in order to maintain homeostasis ^{42,46,47,49}. Tregs employ several mechanisms to suppress the immune response: release of anti-inflammatory cytokines, such as IL-10, IL-35 and transforming growth factor- β (TGF- β), and granzyme A and granzyme B, which induce cell death; metabolic disruption and cytokine deprivation; antigen-presenting cells (APCs) targeting and inhibition of effective antigen presentation to lymphocytes ^{47,49}. Scientific evidence showed that the subsets of Th cells are capable of differentiating into another subset, revealing a certain degree of plasticity ⁴⁶.

Cytotoxic T lymphocytes (CTLs) are cells that express the cluster of differentiation 8 (CD8⁺) and $\alpha\beta$ TCR. Their method of action is direct destruction of cells expressing the target antigen in the major histocompatibility complex I (MHC I) by releasing granules that induce apoptosis or by triggering FasL-mediated cell death. CTLs are the ideal immune cells for cancer cell destruction ^{42,50}.

B cells are another class of lymphocytes that express a B cell receptor (BCR) and can interact directly with antigens or function as APCs in the interaction with T cells. Their main function is the production of antibodies that aid in the immune response ⁴².

Lymphocytes can also be part of innate immunity, as previously described. $\gamma\delta$ T cells are lymphocytes that develop in the thymus and present an $\gamma\delta$ TCR, pursuing their target through small non-peptide antigens (phosphoantigens) that reveal abnormal cellular behavior, such as cancer cells ⁵¹. NKT cells also arise in the thymus and express the $\alpha\beta$ TCR. They primarily act in cytokine production and cell-cell interactions, mimicking T_h1 or T_h2/T_{reg}-like responses against cells and leading to either tumor destruction or escape, respectively ⁵².

Antigens are short sequences of molecules produced by the cell or available in the environment. It acts as a code bar, representing molecules produced by the cell ³⁶. These endogenous antigens are bound to the major histocompatibility complex class I (MHC I) that acts as an anchor for the antigen. APCs are also able to process antigens from the environment and present them to lymphocytes in the major histocompatibility complex II (MHC II), displaying information of potential infection and activating downstream mechanisms ⁴². When antigen-loaded MHC comes into contact with the $\alpha\beta$ TCR match, signals lead to the reprogramming of the naïve T cell in order to start an immune response and stimulate differentiation into activated helper T cells or activated cytotoxic T cells, in an intricate response that results from cooperation from both innate and adaptive immune entities ³⁶.

1.2.1. Tumor cells evade immune surveillance

As previously referred, one of the hallmarks of cancer is the evasion from the immune system ⁶. Although these mechanisms have evolved to effectively protect us against any threats, cancer cells are able to survive the selective pressure and evade immune surveillance, achieving this by either resisting apoptosis, altering immune cells action directly or indirectly, and inducing immunosuppression and tolerance ⁸. The modulation of the immune response to and by cancer cells is denominated immunoediting.

The cancer immunoediting hypothesis recognizes that both innate and adaptive immunity play a key role in cancer prevention but can also model neoplasia and influence disease progression. This hypothesis proposed three phases of the immunoediting process: elimination, equilibrium and escape (**Figure 5**) ⁵³. In the elimination phase, the tumor is completely eradicated by the immune system, which is accomplished by the recruitment of innate immune cells to the neoplastic region and consequent inflammation. This progresses to the point where an interface between innate and adaptive system occurs, as APCs start to survey for unknown antigens specific to the tumor cells and begin activation of T cells, which lead to the successful destruction of all tumor cells ^{41,53}. However, the immunosurveillance process, that occurs in the elimination phase, can function as a Darwinian selective factor, promoting the growth of tumor clones harboring new mutations that reduce immunogenicity ^{40,41}. On other hand, the equilibrium phase is characterized by the constant pressure of the immune system that becomes strong enough to contain the tumor without being able to effectively destroy cancer cells. Transformed cells cycle between dormancy and dividing stages for years, accumulating new mutations and increasing their genetic instability, while continuously being attacked by immune cells. Once

all the right conditions are gathered, tumor cells are then able of overpowering the immune system and progress to the next phase: escape ^{41,53,54}. However, the question that arises here is to elucidate how this shift occurs. Tumor cells need to find niches that can ensure a certain level of protection against the immune system. These niches are composed of endothelial cells, stromal cells, and ECM. Permissive niches are ideal, as triggering angiogenesis can fuel cell divisions and consequent mutations and increased genetic instability that will allow tumor cells to evade the immune system ⁵⁴. Tumor cells are then able to employ a multitude of mechanisms to hide or directly modulate the immune response. These cells also become less immunogenic by decreasing the expression and presentation of MHC class I antigens. Moreover, these tumor cells also directly regulate the immune response by upregulating the expression of checkpoint inhibitors, including programmed death ligand 1 (PD-L1), to suppress T cell activity and increase motility and invasiveness ⁵⁵. In addition, signaling molecules induce recruitment of Tregs and myeloid-derived suppressor cells (MDSC), to the tumor site, leading to secretion of anti-inflammatory cytokines. Lastly, tumor cells overexpress apoptosis-resistance and proliferation genes, in particular BCL-2 and STAT3 ^{40,54–58}.

Together, all of these mechanisms are powerful propulsors of carcinogenesis. Recently, novel research has uncovered new therapeutic approaches exploiting the knowledge accumulate over the years regarding these evasion mechanisms, such as cytotoxic T lymphocyte antigen 4 (CTLA-4) and PD-L1 immunotherapies. These therapies help boost the immune response against cancer and although it can be a promising therapeutic approach to reduce the burden of this disease, a significant percentage of patients often exhibit some degree of resistance ^{20,40,56,58,59}.

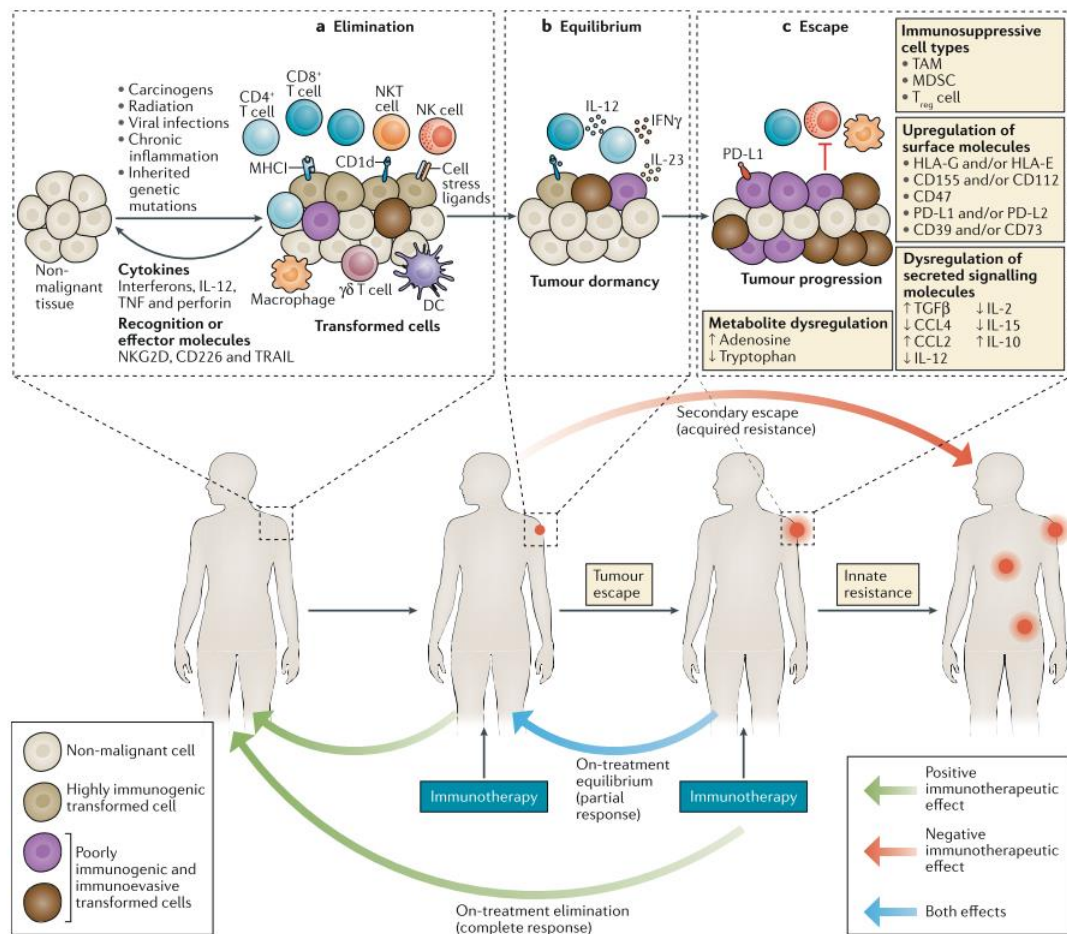


Figure 5 | Cancer immunoediting phases. Tumor cells can be eliminated by the immune system through activation of innate and adaptive response (elimination phase). In the equilibrium phase, tumor cells are kept under pressure by the immune system, cycling between dormancy and division. Once they accumulate the necessary mutations, they can easily escape the immune cells, leading to disease progression (escape phase). Mechanisms employed by tumor cells to suppress the immune system can be exploited to develop novel therapies. From ⁵⁶.

1.3. Glycosylation

Glycosylation consists of the synthesis of sugar links between carbohydrates (also denominated glycans) and other sugars, proteins, and lipids. This is a widely occurring modification observed in several organisms that modulates several cellular mechanisms ^{60,61}. Glyconjugates are classified according to the non-glycosyl component and respective linkages. Glycoproteins encompass proteins with glycosyl linkages such as *N*-linked glycans, *O*-linked glycans, glycosaminoglycans/proteoglycans, and glycosylphosphatidylinositol (GPI) anchors to the backbone, phosphorylated glycans, and C-mannosylation. Glycolipids consist of sugar structures bound to lipids, such as glycosphingolipids (**Figure 6**) ⁶².

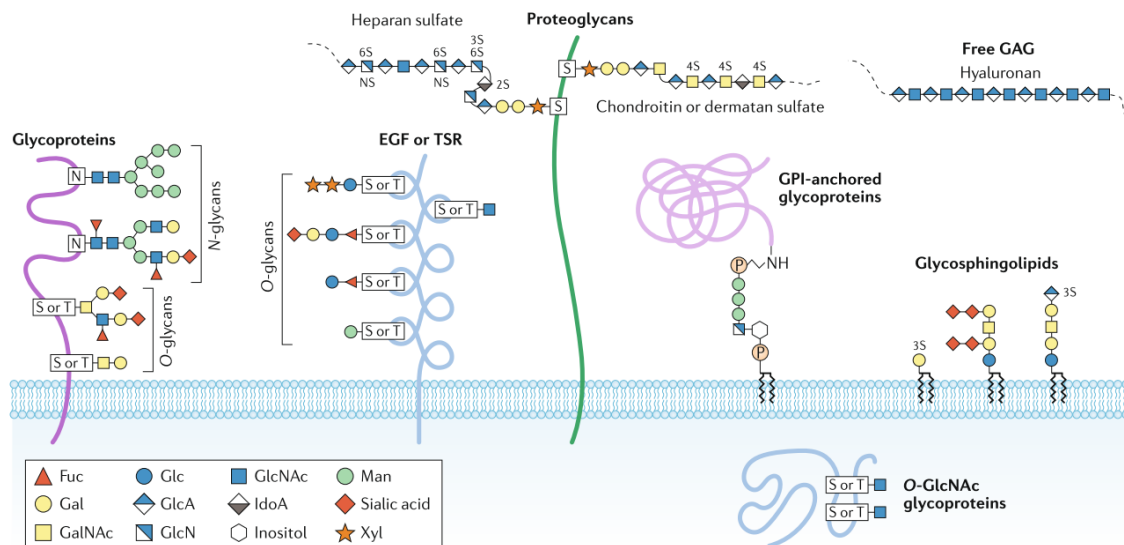


Figure 6 | Different types of glycosylation that occurs in human cells. Glycans are sugar structures found ubiquitously in all life forms. Several distinct types of glycans exist, such as *N*-glycans, *O*-glycans, proteoglycans, glycosphingolipids, glycosylphosphatidylinositol (GPI)-linked proteins. From ⁶².

Glycan biosynthesis occurs predominantly in the cell secretory pathway, more precisely in the endoplasmic reticulum (ER) and Golgi apparatus (GA), involving an orchestra of enzymes that add or remove sugars, denominated glycosyltransferases and glycosidases, respectively ^{62,63}. Nevertheless, it is important to note that some forms of glycosylation, such as GlcNAcylation, occur outside of the secretory pathway ⁶³. Glycosylation is modulated at post-transcriptional and post-translational levels, with glycogene expression, protein activity and location, and substrate availability playing a role in alterations in a cell glycome ^{62–64}. Glycans modulate several cellular processes, including antigen recognition, cell adherence, receptor activation, endocytosis and cell interaction ⁶³. In this sense, it is easily understandable that glycosylation dysregulation plays a crucial role in the pathogenesis of numerous diseases, including IBD ⁶⁵ and cancer ⁶⁰.

N-glycosylation will be addressed in the following section since it is the main focus of this work.

1.3.1. *N*-linked glycosylation

As mentioned above, proteins are frequently glycosylated, as this is the most frequent and diverse post-translational modification that occurs in these molecules ⁶⁴, with around 70% of proteins being glycosylated ⁶⁶. More or less 90% of these glycoproteins are characterized by the presence of *N*-linked glycans, structures consisting of a common core

of 2 *N*-acetylglucosamine and 3 mannose residues (GlcNAc₂Man₃), followed by a variable number and order of monosaccharides ^{62,67}. The core structure is bound to the protein by a link to the nitrogen atom of specific asparagine residues in a consensus sequence, defined by asparagine - any amino acid, except proline - serine or threonine (Asn-X-Ser/Thr) (**Figure 7**). In eukaryotic cells, the process begins with the addition of *N*-Acetylglucosamine-phosphate (GlcNAc-P) to the dolichol-phosphate (Dol-P) lipid carrier in the membrane of the ER, a process catalyzed by cytoplasmic-facing membrane-embedded glycosyltransferases (GTs). This GlcNAc-P is derived from a uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) ⁶⁷. Then, another *N*-acetylglucosamine (GlcNAc) residue and five mannose (Man) residues are added to the structure. Next, the structure is flipped to the lumen of the ER, where 4 more Man residues are added, finalizing with 3 glucose (Glc) residues. Once this glycan precursor is formed, the oligosaccharyltransferase (OST) enzyme transfers it *en bloc* to the nascent protein. After this 14-sugar precursor is attached to the protein at the Asn-X-Ser/Thr, trimming and removal of glucose residues are carried out. Finally, glycoproteins pass through quality control and are then allowed to proceed to the Golgi apparatus for further processing ⁶⁸. In *cis*-Golgi, glycans are trimmed once again by the action of α -1-2-mannosidases IA, IB, and IC, giving rise to Man₅GlcNAc₂, an important intermediate structure in the pathway that produced hybrid and complex *N*-glycans ⁶⁹. *N*-acetylglucosaminyltransferase I (GnT-I) adds a GlcNAc residue to the α -1,3-linked Man and two Man residues are removed by α -mannosidase II. Moving to the *medial*-Golgi, GnT-II then adds another GlcNAc to the α 1-6Man, leading to biantennary *N*-glycans. Further branching can occur by action of *N*-acetylglucosaminyltransferase IV (GnT-IV) (addition of a GlcNAc residue the α -1,3-linked Man via a β -1,4-linkage) and/or *N*-acetylglucosaminyltransferase V (GnT-V) (addition of a GlcNAc residue the α -1,6 Man via a β -linkage), while bisecting branches can be formed by GnT-III ⁷⁰ (**Figure 7**).

Glycans mature in the *trans*-Golgi, where the final modifications are made. Fucose residues are added in a α -1,6-linkage to the core GlcNAc, branches can be elongated with poly-*N*-acetylglucosamine (poly-LacNAc) and caps (such as fucose, sialic acid, galactose) (**Figure 7**) ⁷⁰.

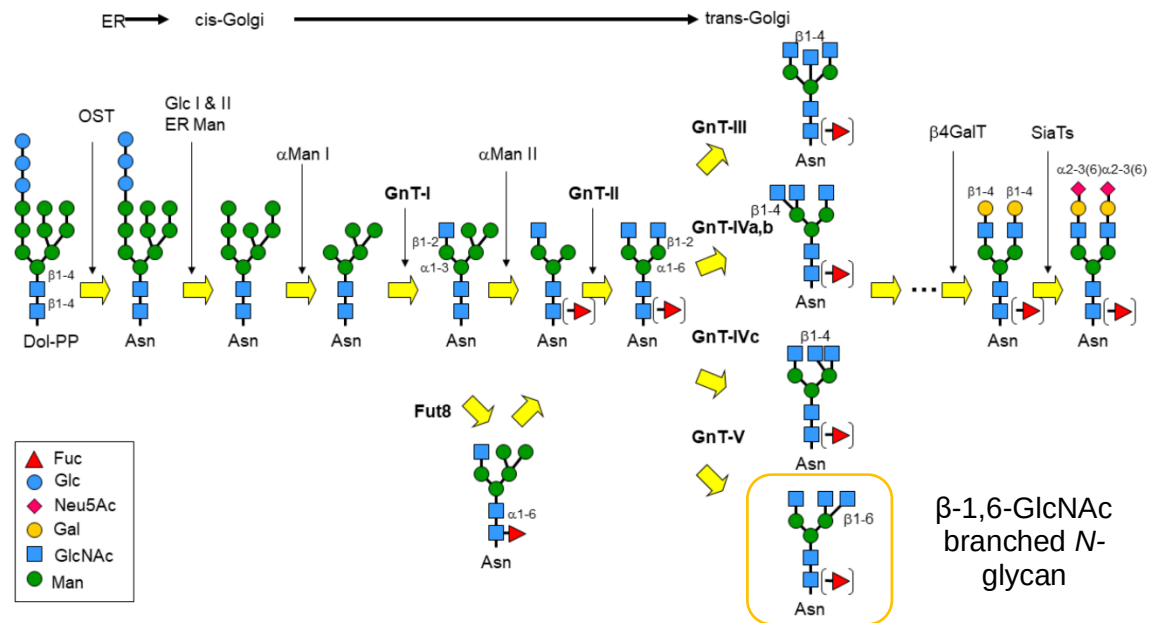


Figure 7 | Simplified N-glycan biosynthesis pathway on the secretory pathway. Glycans are subject to the action of several enzymes, which add and remove sugar residues to the core structure. The pathway also relies on the action of important transport proteins. N-glycans can further be classified as oligomannose, hybrid or complex. Adapted from ⁶¹.

1.3.2. The glycosylation impact on T cell activity and disease

Glycosylation directly modulates T lymphocyte development and activity. Firstly, glycans play an important role in homing of thymus-seeding progenitors (TSPs) originating from the bone marrow. P-selectin is a glycoprotein expressed in thymus epithelial cells that allows the homing of T lymphocytes expressing P-selectin glycoprotein ligand-1 (PSGL-1). The next critical steps of T cell maturation and positive and negative selection are equally intimately regulated by the glycan structures ⁷¹. Moreover, glycosylation also regulates several T cells functions, such as differentiation, signaling, self-renewal, and can even be responsible for immune tolerance and suppression (**Figure 8**). These mechanisms are modulated by the interaction of glycan-binding proteins (GBPs) with the glycan structure ⁷².

Recent research has revealed how glycans play a role in the pathophysiology of several diseases ^{62,63,73}. Interestingly, our group found that lamina propria T lymphocytes from UC patients exhibited a deficiency in complex branched β-1,6-N-glycans on the TCR, consequence of a decreased expression of the *α-1,6-mannosylglycoprotein 6-β-N-Acetylglucosaminyltransferase A (MGAT5)* gene ⁶⁵. This profile was correlated with a decreased T cell activation threshold, leading to an excessive inflammatory response ⁶⁵.

Reversal of this glycoprofile was achieved by metabolic supplementation with GlcNAc, controlling inflammation and decreasing the severity of the disease ⁷⁴.

On the other hand, glycosylation is also implicated in cancer. Dysregulation of glycosylation, by either overexpression or underexpression of glycosyltransferases and glycosidases, leads to the synthesis of distinct glycan structures, but can also be responsible for carbohydrate neo-synthesis (*de novo* expression of determined antigens) ⁶⁰. Alterations that are often observed include an increase in sialyl Lewis structures, aberrant core fucosylation, and also abnormal *N*-glycan branching, all of which contribute to a pro-malignant phenotype ⁶². In fact, β -1,6-GlcNAc branched *N*-glycans are overexpressed in gastric cancer cells, namely in E-cadherin, leading to an inhibition of its normal activity and promoting cancer cell motility and invasiveness ^{75–77}. In addition, to contributing with new selective advantages for tumor cells to grow and spread, glycosylation in these cells can also be employed as a mechanism for immune suppression ⁷². Indeed, our group recently demonstrated that the same overexpression of these branched *N*-glycans occurs in CRC cells and participates in an immunosuppressive mechanism in which transformed cells mask themselves from host defense systems, decreasing antigen presentation and repressing immune recognition by GBPs, internalizing MHC I molecules and inhibiting IFN- γ production ⁷⁸. The glycoreprogramming of these cells with kifunensine (KF), an inhibitor of α -mannosidase I, leads to decreased expression of branched *N*-glycans in favor of high-mannose structures, which trigger the immune response and tumor suppression ⁷⁸.

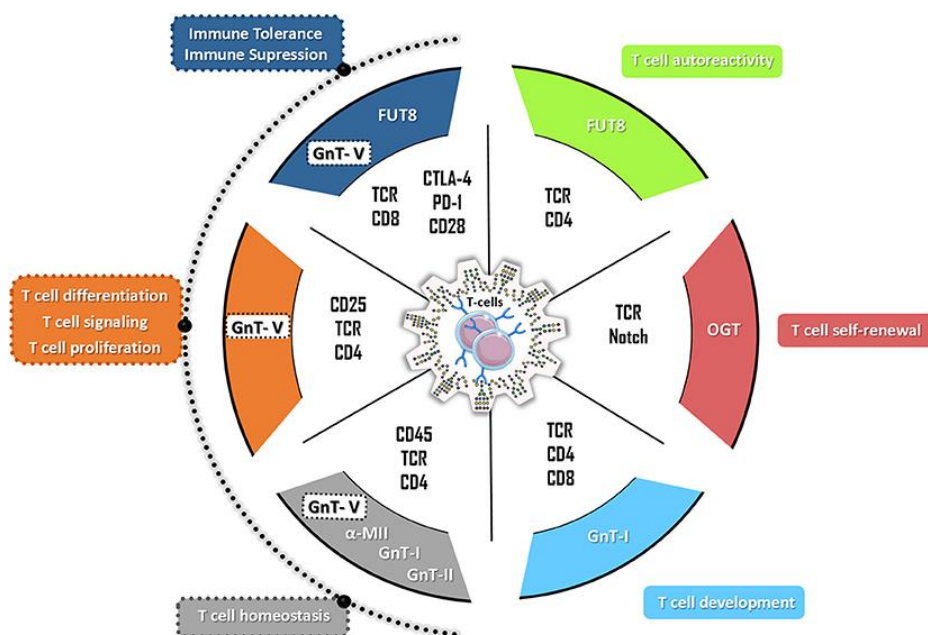


Figure 8 | Glycans modulate several T cell mechanisms. The complex process of glycosylation is a key regulator of T cell homing to the thymus, development and selection, differentiation, proliferation, and homeostasis. Glycans also participate in immune tolerance, suppression, and autoreactivity observed in several diseases. From

⁷¹.

Taken together, this evidence highlights glycans as promising novel therapeutic targets for colorectal cancer treatment ⁷². For this purpose, it is absolutely vital to not only understand the mechanisms involved in cellular dysregulation in cancer, but also uncover how differential glycosylation impacts cancer immunity at the immune-tumor interface. In fact, there is a major gap in knowledge regarding the glycoprofile of immune cells along the carcinogenic pathway and tackling this issue is essential in order to decipher how glycan dysregulation leads to an unresponsive immune system and allows cancer cells to thrive

2. General aim

The main goal of this thesis is to unravel the impact of *N*-glycosylation in the T lymphocyte mediated immune response in distinct colorectal cancer (CRC) pathways and its precursor lesions. Given the enormous importance of the glycome cancer development and progression and particularly in the modulation of the immune system, we herein aim to uncover novel mechanistic insights in T cell *N*-glycosylation and how this alters the specific response against pre-malignant lesions and CRC. Knowledge regarding the changes in T cell glycoprofile along CRC cascade can promote the development of novel preventive therapeutic approaches by cellular glycoreprogramming. Therefore, three major aims were defined:

1. To investigate the changes in T cells glycoprofile in pre-malignant and cancer lesions of distinct CRC cascades and to assess the impact of the altered expression of β -1,6-branched *N*-glycans in T cells in the development and progression of CRC;
2. To assess how β -1,6-branched *N*-glycan expression affects immune cell activity and interaction with cancer cells;
3. To evaluate the potential of T cell glycoreprogramming as a useful preventive approach for CRC.

To achieve this, we have extensively studied the glycoimmune profile of lesions from serrated pathway, Lynch syndrome and colitis-associated colorectal (CAC) cancer pathways, using human formalin-fixed paraffin-embedded (FFPE) samples for both glycans and protein expression analysis by lectin and immuno histochemistry, as well as gene expression analysis by real time- quantitative polymerase chain reaction (RT-qPCR). We also performed lectin histochemistry on FFPE samples from wild type and *Mgat5*^{-/-} mice (which were subjected to induced colitis and carcinogenesis by azoxymethane and dextran sodium sulfate (AOM/DSS)) to further validate our human results. Moreover, fresh biopsies and blood were collected from serrated and CAC cascade cases in collaboration with Centro Hospital Universitário do Porto (CHUP) hospital. Flow cytometry analysis was performed to further quantify immune populations infiltrating into lesions and study their glycoprofile. Furthermore, we studied how glycans modulate cell interaction at the immune-tumor interface by glycoreprogramming mice T lymphocytes to either under-express or overexpress β -1,6-branched *N*-glycans. We then performed successive cocultures assays with a murine CRC cell line and analyzed tumor cell viability by flow cytometry, allowing us to evaluate the differential T cell response against cancer cells based on the specific glycoprofile.

3. Materials and Methods

3.1. Characterization of patient cohorts from CAC and CRC pathways

The present work includes the use of biological material from patients diagnosed with colitis-associated colorectal cancer (CAC) cascade lesions (this cohort will henceforth be denominated as IBDCan cohort) and serrated CRC and Lynch syndrome cascades (this cohort will henceforth be denominated as pre-malignant cohort). These patients were admitted in the Gastroenterology Department of Centro Hospitalar Universitário do Porto (CHUP), where these samples were obtained. Patient selection and recruitment were performed by Dr. Paula Lago, Dr. Ricardo Marcos-Pinto and Dr. Mónica Garrido. Each individual patient signed an informed consent to donate samples for the study. The IBDCan cohort also includes biological samples from the Hospital Clinico San Carlos – Madrid (international collaboration with Dr. Carlos Taxonera).

The IBDCan project cohort included 71 formalin-fixed paraffin embedded (FFPE) samples from colitis patients, which were subdivided into categories according to the risk of developing CAC (according to established ECCO guidelines ⁷⁹). Severity of inflammation/colitis is classified endoscopically and ranges from Mayo 1 to 3. **The low risk of developing CAC** (n=26) includes patients with extensive colitis with mild active endoscopic or histological inflammation (Mayo 1) and more than 8 years of diagnosis. **Intermediate risk of CAC** (n=22) includes patients with extensive colitis with moderate or severe endoscopic or histological inflammation (Mayo 2 and 3), history of CRC in a first degree relative older than 50 years, and those presenting post-inflammatory polyps (hyperplastic and pseudo polyps) and more than 8 years of diagnosis. **High-risk** (n=7) patients are those with pancolitis with moderate or severe inflammation and more than 8 years of diagnosis. This cohort also included patients with **low grade dysplasia (LGD)** (n=3) and **colorectal carcinoma** (n=1). No patients with high grade dysplasia (HGD) were included. In addition, n=7 FFPE samples from normal adjacent colonic mucosa/healthy controls and n=26 samples from remission/adjacent mucosa were included in the study. These samples were analyzed by lectin and immunohistochemistry, as well as by real time-quantitative polymerase chain reaction (RT-qPCR).

In addition to the FFPE samples used, fresh colon biopsies and blood were collected at CHUP (n=2 patients) and used for fluorescence-activated cell sorting (FACS) analysis. Fresh colonic biopsies from each patient included, whenever possible, a sample of the lesion and a sample of adjacent mucosa. The histopathological examination and classification of fresh biopsies were performed by the Department of Pathology in CHUP.

The pre-malignant lesions cohort included FFPE samples from patients diagnosed with precursor lesions or CRC of the serrated (n=33) and Lynch pathways (n=11). Regarding the serrated pathway, samples were subdivided into the following categories: **sessile serrated adenoma without dysplasia** (n=22), **sessile serrated adenoma with dysplasia** (n=8) and **serrated adenocarcinoma** (n=3). Lynch pathway samples were subdivided into **low grade dysplasia (LGD)** (n=6), **high grade dysplasia (HGD)** (n=2) or **Lynch CRC** (n=3). In addition to these samples, 7 normal adjacent colonic mucosa/healthy control FFPE samples were included. These samples were analyzed by lectin and immunohistochemistry in order to study glycan and protein expression. Furthermore, messenger RNA (mRNA) was extracted, and gene expression was analyzed by RT-qPCR.

This cohort also included fresh colonic biopsies collected in CHUP from patients with serrated sessile adenomas without dysplasia lesions (n=8), patients with serrated sessile adenomas with dysplasia lesions (n=2) and patients with colon adenocarcinoma (n=2). In addition, normal adjacent mucosa/healthy colonic samples were included as controls (n=6). These samples were processed and analyzed by FACS to study the specific glycoimmune profile. Once again, the histopathological evaluation was performed by pathologists at CHUP.

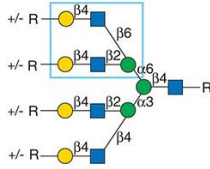
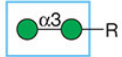
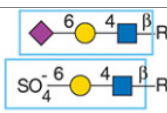
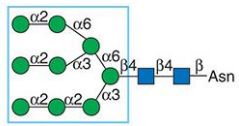
3.2. Lectin histochemistry

Lectins are glycan-binding proteins (GBPs) extensively used in research that recognize specific domains, being useful in the discrimination of glycan structures (**Table 1**)⁸⁰.

FFPE sample blocks were sectioned at 3 μ m using a microtome (paraffin microtome Microm HM335E) and placed in slides. These slides were then deparaffinized with xylene and hydrate using alcohol solutions with decreasing concentrations (twice in alcohol 100 % and once in alcohol 70%) and washed with tap water. Next, they were placed in a 3% hydrogen peroxide solution for 10 minutes to inactivate endogenous peroxidases. The slides were then washed with 1x phosphate buffered solution (PBS) and the sections were delimited with a hydrophobic pen. Protein blocking was done in a humid chamber with either bovine serum albumin (BSA) 10% or Carbo-free Blocking solution (Vector Laboratories) 1x diluted in PBS 1x for 30 minutes. Subsequently, slides were incubated with the desired biotinylated lectin diluted appropriately in PBS 1X for 1 hour: *Galanthus nivalis* (GNA) lectin (1:100 dilution; Vector Laboratories), which recognizes terminal α 1-3-linked mannose, or *Phaseolus vulgaris* Leucoagglutinin (L-PHA) lectin (1:100 dilution; Vector Laboratories), which recognizes β -1,6-GlcNAc complex branched N-glycans. After lectin incubation, slides were washed with PBS 1X and incubated 30 minutes at room temperature (RT) with the

avidin-biotin complex (ABC) solution (dilution 1:100 in PBS 1X; Vectastain® ABC Kit, Vector Laboratories). After being washed once again with PBS 1X, slides were stained for 2-3 minutes with diaminobenzidine (DAB) solution (ThermoFisher Scientific). Slides were then washed with running water, counterstained with hematoxylin, washed once again with tap water and then dehydrated with alcohol solutions with increasing concentrations of ethanol, finalizing in xylene. To conclude the process, slides were then mounted on Entellan® mounting medium, left to air dry and observed on a Zeiss Axioskop 2 microscope. Images representative of the region of interest were captured with a Nikon DS-L1-2M camera. Slides were then quantified based on the intensity of signal: 0 if no signal was present; 1 if a weak signal was observed; 2 for an intermediate intensity signal; and 3 for a strong signal.

Table 1 | Characterization of lectins used in the present work. Lectins are glycan-binding proteins (GBPs) that recognize specific glycan domains. Glycans structures adapted from chapter 48 of the Essentials of Glycobiology⁸⁰.

Lectin abbreviation	Lectin denomination	Glycan binding domain
L-PHA	Phytohemagglutinin-L from <i>Phaseolus vulgaris</i>	 Recognizes β-1,6-GlcNAc
GNA	<i>Galanthus nivalis</i> agglutinin	 Recognizes α-1,3-Man
SNA	<i>Sambucus nigra</i> agglutinin	 Recognizes terminal α-2,6-linked sialic acid
ConA	Concanavalin A lectin from <i>Canavalia ensiformis</i>	 Recognizes mannose residues

3.3. Immunohistochemistry

FFPE samples followed the same steps for deparaffinization, and hydration referred in section 2.2. After this step, slides were inserted into cuvettes containing 0,05% Extran solution (for CD3 and CD8 immunohistochemistry (IHC)) or citrate buffer (for Foxp3 IHC) and placed on a “Ufesa steam cooker” for 40/45 minutes in order to retrieve tissue antigens. Slides were then left to cool at RT for 20 minutes and washed twice with PBS 1x.

Slides subjected to either CD3 or CD8 IHC were laid in a 3% hydrogen peroxide solution for 10 minutes in order to block endogenous peroxidases. Afterwards slides were washed twice with PBS 1x, and sections were delimited with the hydrophobic pen. Tissue was then incubated with swine normal serum (in case of CD3 IHC; dilution 1:5 in BSA 10%; Dako) or rabbit normal serum (in case of CD8 IHC; dilution 1:5 in BSA 10%; Dako) for 30 minutes at RT. After this, antibodies were placed onto sections: rabbit anti-human monoclonal CD3 (dilution 1:50 in BSA 5%; clone SP7 from ThermoFisher Scientific) or mouse anti-human monoclonal CD8 (dilution 1:50 in BSA 5%; clone 4B11 from ThermoFisher Scientific) and slides were left incubating overnight at 4°C. In the next day, slides were washed with PBS 1x twice and incubated with swine anti-rabbit biotinylated polyclonal antibody (for CD3 IHC; dilution 1:200 in BSA 5%; Dako) or rabbit anti-mouse biotinylated polyclonal antibody (for CD8 IHC; dilution 1:200 in BSA 5%; Dako). After another round of washes with PBS 1x, slides were incubated with the ABC solution for 30 minutes at RT, washed once again and revealed with DAB solution for 5/7 minutes. The next steps were the same as described in the lectin histochemistry protocol, including hematoxylin staining, dehydration, diaphanization and mounting.

Slides that were subjected to Foxp3 IHC had their sections delimited with the hydrophobic pen and incubated with Endogenous Peroxidase Blocking Solution (UltraVision™ Quanto Detection System HRP kit from ThermoFisher Scientific) for 10 minutes. Afterwards, they were washed with PBS 1x and incubated for 8 minutes with protein block solution (UltraVision™ Protein Block from ThermoFisher Scientific). Finally they were left overnight at 4°C incubating with the mouse anti-human Foxp3 antibody (dilution 1:100 in BSA 5%). In the next day, slides were washed with a PBS 1x + 0,01% Tween (PBST) solution, incubated for 10 minutes with Primary Antibody Amplifier Quanto and washed once again. Slides were then placed in the humid chamber protected from light while incubation with HRP Polymer Quanto occurred (10 minutes). After washing with 0,01% PBST, slides followed the previously described protocol regarding revelation, counterstaining, dehydration, diaphanization and mounting. All of the slides were then observed on a Zeiss Axioskop 2 microscope and evaluated. Images representative of these region of interest were captured with a Nikon DS-L1-2M camera. Slides were then quantified based on the intensity of signal: 0 if no signal was present; 1 if a weak signal was observed; 2 for an intermediate intensity signal; and 3 for a strong signal.

3.4. RNA extraction from FFPE samples

FFPE blocks were sectioned at 10 µm using the paraffin microtome Microm HM335E and samples were placed into eppendorfs. RNA extraction was performed following the

Recover All™ Total Nucleic Acid Isolation Procedure Kit (ThermoFisher Scientific) manufacturer protocol. Samples were deparaffinized in xylene for 3 minutes at 50 °C to remove the paraffin from tissues. The eppendorfs were then centrifuged for 2 minutes at maximum velocity and the supernatant was discarded. The pellet was washed twice with 100% ethanol, finalizing with the samples left to air dry until all ethanol evaporates (approximately 2 hours at RT). Samples were then incubated with digestion buffer and protease mix for 15 minutes at 50 °C and 15 minutes at 80 °C. Afterwards, samples were incubated with isolation additive/ethanol mix and placed on a separation column. After several washes (with Wash 1 and Wash 2/3) and centrifugations (for 30 seconds at 10,000 rpm), DNase mix was placed on the columns to destroy DNA contaminants, and the samples were washed twice again. For the final step, the elution solution was placed on the column filter and the samples were centrifuged, concluding with the collection of RNA into the eppendorfs. Extracted RNA was then quantified in a NanoDrop 1000 UV-Vis Spectrophotometer system (ThermoFisher Scientific) and stored at -80 °C until further use.

3.5. cDNA synthesis and real-time quantitative polymerase chain reaction (RT-qPCR)

FFPE extracted RNA was transcribed into cDNA using the SuperScript™ IV Reverse Transcriptase kit (ThermoFisher Scientific). RNA was diluted in ultra-pure nuclease-free water and incubated with deoxyribonucleotide triphosphate (dNTPs; 1 mM) and random primers (0,1 µg/µl). The samples were then placed on the T100™ thermocycler (Bio-Rad Laboratories) with a set of two program cycles: 5 minutes at 65°C and infinite hold at 4°C. A reaction mix including buffer solution, DTT, RNase Out enzyme and reverse transcriptase enzyme was then added, and the samples were once again placed on the thermocycler for another program: 10 minutes at 23°C, followed by 10 minutes at 52°C, 10 minutes at 80°C and finally infinite hold at 4°C. The cDNA samples were stored at -20°C until further use.

RT-qPCR was performed on 96-well reaction plates (Applied Biosystems by Life Technologies). First, a reaction mix, consisting of 5 µl of TaqMan™ Universal PCR Master Mix (ThermoFisher Scientific), 2,5 µl of nuclease free water and 0,5 µl of the required probe, was prepared (values of reaction mix for 1 well). The list of probes utilized in this work can be found in **Table 2**. When performing multiplex RT-qPCR (two amplification probes in the same well), the reaction mix consisted of 5 µl of TaqMan™ Universal PCR Master Mix, 2 µl of nuclease-free water, 0,5 µl of probe_1 and 0,5 µl of probe_2. Finally, 2 µl of cDNA was added to each well. The template was designed using 7500 Software version 2.0.6, including the identification of wells with the respective sample and probe. The 96-well plate

was then placed in a 7500 real-time PCR system (Applied Biosystems) with a standard ramp speed and the following settings: holding stage at 50°C for 20 seconds and 95°C for 10 minutes; and 45 amplification cycles consisting of 15 seconds at 95°C and 1 minute at 60°C. Results from gene expression were normalized for *18S* or *CD3ε* and relative quantification (RQ) values were calculated following the $2^{-\Delta CT}$ procedure.

Table 2 | Amplification probes used for gene expression quantification by RT-qPCR. All probes used are from TaqMan™, except the *18S* probe which is manufactured by Integrated DNA Technologies (IDT).

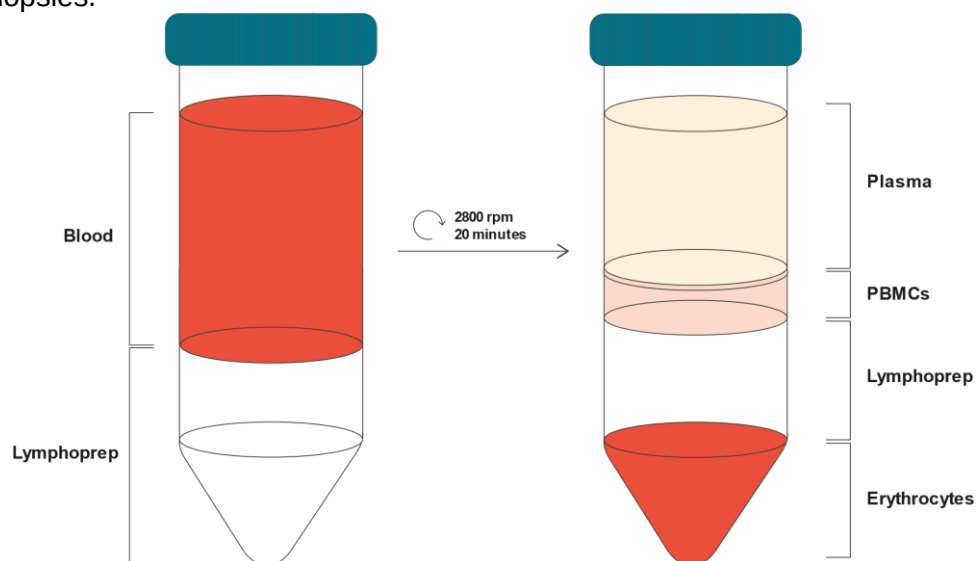
Gene	Gene designation	Probe identification	Reporter dye
MGAT5	<i>α-1,6-Mannosylglycoprotein 6-β N-Acetylglucosaminyltransferase</i>	Hs00159136_m1	FAM
MAN2A1	<i>Mannosidase α class 2A member 1</i>	Hs01123597_m1	FAM
ST6GAL1	<i>ST6β-galactoside α-2,6-sialyltransferase 1</i>	Hs00949382_m1	VIC
FUT8	<i>Fucosyltransferase 8</i>	Hs001896535_m1	VIC
FOXP3	<i>Forkhead Box P3</i>	Hs01085834_m1	FAM
CD8α	<i>T-cell surface glycoprotein CD8 alpha chain</i>	Hs00233520_m1	FAM
CD3ε	<i>T-cell surface glycoprotein CD3 epsilon chain</i>	Hs01062241_m1	FAM
18S	<i>18S ribosomal RNA</i>	Custom_Human (designed by P.Oliveira)	FAM

3.6. Fresh colonic biopsy and blood processing and FACS analysis

Fresh biopsies of premalignant lesions from the serrated pathway and of CAC cascade lesions/IBDCan cohort were processed for FACS analysis. This process involved mechanical disruption of the tissue in Hanks' Balanced Salt Solution (HBSS) (Sigma) supplemented with 1% penicillin/streptomycin and 0,1% gentamicin, followed by enzymatic digestion of the biopsies in complete Roswell Park Memorial Institute (RPMI) medium (Gibco® from ThermoFisher Scientific) with collagenase using an orbital shaker for 45 minutes at 37°C. The resulting suspension was then forced through a cell strainer to obtain a final single cell suspension, centrifuged, and resuspended in PBS 1x. The suspension was then pipetted into the 96-well plate in distinct wells (for the different lectin and antibody mixes) and centrifuged once again to discard PBS 1x.

At this point, the FACS staining was initialized, starting with incubation with the fixable viability dye (FVD, 1:2000 dilution) for 20 minutes on ice and protected from light. After this incubation, the cells were washed with PBS 1x and centrifuged at 2000 rpm for 1 minute. The supernatant was discarded. The next step involved incubation with lectins. The lectins utilized in this process were L-PHA, GNA and *Sambucus nigra* agglutinin (SNA), which recognize distinct glycan structures of interest (**Table 1**). Samples were left incubating for 15 minutes on ice and protected from light and afterwards were washed with PBS 1x and centrifuged. The samples were then incubated with mice serum 2% for 10 minutes on ice and protected from light, followed by incubation with the respective mix of primary antibodies for 30 minutes on ice and protected from light (**Table 3**). Cells were then washed with FACS buffer (PBS 1x and 1% fetal bovine serum (FBS) (Biowest) and centrifuged. The samples were then incubated with secondary antibody (Streptavidin) for 20 minutes on ice and protected from light, washed with FACS buffer, and centrifuged. Cells were fixed with 2% formaldehyde for 30 minutes at room temperature and protected from light or incubated with fix/perm buffer for 30 minutes at room temperature and protected from light (this buffer will permeabilize the cell membrane and allow intracellular staining). Next, cells were centrifuged resuspended in FACS buffer and recovered in cytometry tubes or stained with 25 µl of intracellular antibody (Foxp3) for 30 minutes, at RT and protected from light. Afterwards, Perm buffer was added to the wells and samples were centrifuged. Finally, FACS buffer was added, and samples were collected into cytometry tubes.

Blood samples received from CAC cascade patients were centrifuged with Lymphoprep™ (2800 rpm for 20 minutes at 23°C) to form a gradient separation of serum, peripheral blood mononuclear cells (PBMCs) and erythrocytes (**Figure 9**). The PBMCs interface was recovered and followed the same FACS staining protocol as described for fresh biopsies.



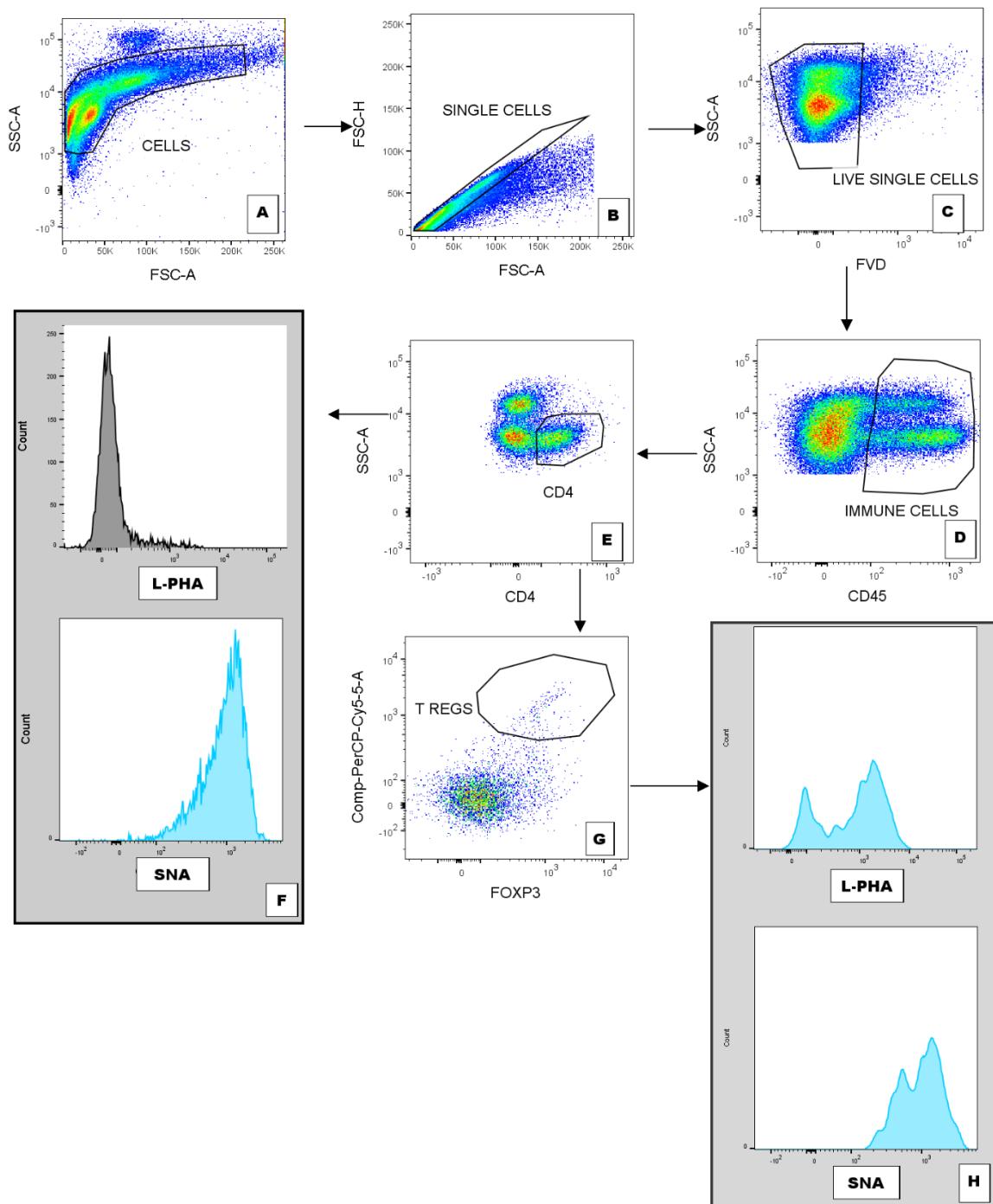
(figure legend in the next page)

Figure 9 | Blood components gradient separation with Lymphoprep™. Lymphoprep solution was pipetted into a falcon followed by the addition of peripheral blood without mixture, forming two layers. The sample was centrifuged at 2800 rpm for 20 minutes at 23°C. Afterwards, 4 distinctive layers could be distinguished based on the density of the components: an erythrocyte layer on the bottom, followed by the Lymphoprep layer, a layer containing the peripheral blood mononuclear cells (PBMCs) and the plasma layer.

Table 3 | Antibodies and lectins used for human samples FACS analysis, with respective fluorochrome, dilution and manufacturer.

Antibody/Lectin	Color channel	Fluorochrome	Clone	Manufacturer	Dilution
L-PHA	FITC	FITC	-	Vector Laboratories	1:1000
GNA		FITC	-	Vector Laboratories	1:1000
CD45	AmCyan	Brilliant Violet 510	HI30	BioLegend	1:200
CD4	Pacific Blue	eFluor450	RPA-T4	eBioscience	1:100
CD3		PB	OKT3	BioLegend	1:100
FVD	APC-Cy7	APC-Cy7	-	eBioscience	1:2000
SNA	APC	APC	-	Vector Laboratories	1:1000
CD25		APC	BC96	BioLegend	1:100
CD4	PercP	PercP-Cy5.5	RPA-T4	BioLegend	1:200
CD25		PercP-Cy5.5	BC96	BioLegend	1:100
Gamma/delta TCR		PerCP-eFluor710	B1.1	eBioscience	1:100
CD8	PE-Cy7	PE-Cy7	RPA-T8	BD Pharmingen	1:400
Streptavidin		PE-Cy7	-	Invitrogen	1:200
Streptavidin	PE	PE	-	BioLegend	1:200
Foxp3		PE	PCH101	eBioscience	1:100
ConA	-	No fluorochrome; biotinylated	-	Vector Laboratories	1:1000
SNA	-	No fluorochrome; biotinylated	-	Vector Laboratories	1:1000

Fluorescence values were then obtained in a FACSCanto II flow cytometer (BD Biosciences) and results were analyzed in FlowJo version 10 software. The analysis process included the percentages of populations of interest and lectin or antibody median fluorescence intensity (MFI) values. For this purpose we employed a gating strategy consisting of selection of viable/live single cells, followed by selection of an immune cell population of interest and research on the specific glycoprofile (**Figure 10**).



(figure legend in the next page)

Figure 10 | General cytometry gating strategy for human biopsies and blood. The analysis started with the selection of cells based on the relation between size (forward scatter or FSC) and complexity/granularity (side scatter or SSC), as demonstrated by **A**. Afterwards, doublets cells were excluded by plotting event height (forward scatter height or FSC-H) by event area (forward scatter area or FSC-A) (**B**). Then live single cells negative for fixable viability dye (FVD) were selected (**C**). As the biopsies are composed of immune populations but also epithelial cells, cells expressing CD45 immune marker were selected (**D**). These populations were then divided on the several immune populations based on the expression of determined immune markers. As demonstrated in this example (**E**), helper T cells were selected by gating cells positive for CD4 immune marker. In this population, regulatory T cells were further selected by gating cells positive for FoxP3 immune marker (**G**). The glycoprofile was then determined by analyzing lectin-glycan interactions (**F** and **H**).

3.7. Cell lines and cell culture conditions

The murine colon adenocarcinoma cell line 38 (henceforth denominated as MC38) were used for cocultures with wild type mice splenocytes in order to study the impact of glycosylation in T cell function. The cell culture conditions for this specific cell line are described below in an x / y format (in which x is the volume of the reagent when using a T25 flask and y is the volume when using a T75 flask).

The frozen MC38 cells were thawed using prewarmed Dulbecco's modified Eagle's minimal essential medium (DMEM; Gibco® from ThermoFisher Scientific) to resuspend the pellet. Thawed cells were transferred to culture flasks with 4/12 ml of DMEM medium supplemented with 20% FBS and 1% of Penicillin/Streptomycin (Pen/Strep) (Gibco® from ThermoFisher Scientific) and placed in an incubator at 37°C and with 5%CO₂. Cells reaching 80% of confluence were subcultured to maintain the cell line. All the DMEM medium used was removed from the flask and 1/3 ml of prewarmed PBS 1x was added to clean the cell monolayer. Afterwards, 1/3 ml of 0,05 trypsin solution was added and flasks were placed in the incubator for 1 minute. After this time, 2/6 ml of fresh DMEM supplemented with 10% FBS and 1% Pen/Strep was added in order to stop trypsin action. Some of the resulting cell suspension was then transferred to a new T25/T75 flask containing 4/12 ml of new DMEM supplemented medium.

3.8. Splenocytes supplementation with kifunensine and *N*-acetylglucosamine

Fresh spleen from a wild type mouse was digested in order to obtain fresh splenocytes. Briefly, the spleen was crushed in RPMI medium in a 6-well plate. Suspension was then forced through a cell strained in order to obtain only single cell suspension. This suspension was then centrifuged at 1200 rpm for 5 minutes at 4°C, supernatant was discarded, and

pellet was incubated with ACK 1x solution (ACK 10x solution diluted in deionized water) for 2 minutes to lyse erythrocytes. Medium was then added to stop the lysis process and samples were once again centrifuged. Then 2 ml of fresh medium was added and splenocytes were counted with trypan blue (1:100 dilution) on a Neubauer chamber. Viable cells counted on the four corner squares were considered and viable cells in total suspension was calculated following the formula:

$$\text{Total splenocytes in suspension} = \frac{\text{Viable cells counted}}{\text{Number of squares counted}} \times \text{dilution factor} \times 10^4 \times \text{suspension volume}$$

After counting cells the determined volume with the required number of cells was pipetted to a new falcon and total volume of RPMI medium needed (200 µl per well) was added. After this time, 96-well plates were coated with anti-CD3 (0,2 µg/µl; diluted in PBS 1x) for 2 hours at 37°C. Then, wells were washed one time with PBS 1x solution. splenocytes were then recovered, centrifuged, and resuspended in fresh medium and placed on the new plate. Cultured cells were treated with either 20 µM of kifunensine (KF) or 100 mM of *N*-acetylglucosamine (GlcNAc) (**Figure 11**) for 24 hours; 48 hours; and for 48 hours with supplementation and 24 hours without (concentrations already described in the group). After each timepoint, cells were stained with FVD (dilution 1:2000) for 20 minutes (on ice and protected from light) and fixed with PFA 2% (protected from light at RT). Finally, all timepoint plates were stained for lectins L-PHA, GNA, SNA and Concanavalin A lectin from *Canavalia ensiformis* (ConA) (25 µl of mix for 15 minutes, on ice and protected from light). Then, samples were stained with 25µl of primary antibodies (anti-CD3; dilution 1:100) (30 minutes on ice and protected from light), washed, and incubated with secondary antibodies (streptavidin-PE) (30 µl, 20 minutes on ice). Samples were then acquired on FACS Canto flow cytometer and analyzed on FlowJo version 10 software in order to determine the best timepoints for each supplementation.

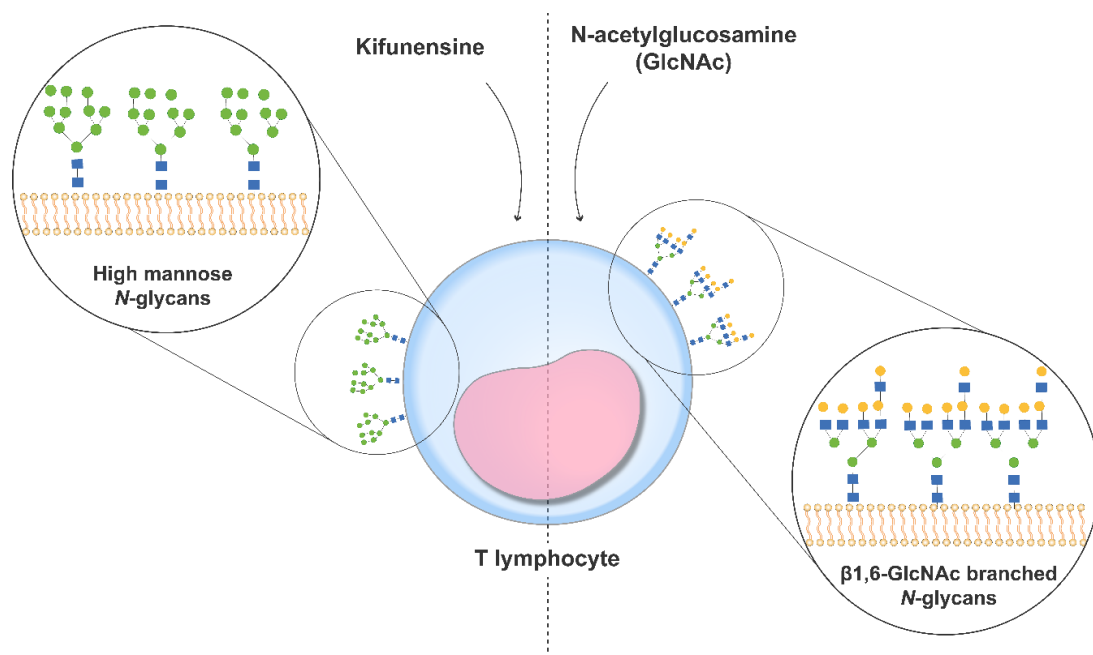


Figure 11 | Effects of kifunensine (KF) and N-acetylglucosamine (GlcNAc) supplementation in T cell glycome. Kifunensine (KF) is a powerful inhibitor of the α -mannosidase I enzyme, blocking the *N*-glycosylation pathway to form more complex *N*-glycans and promoting high-mannose structures. *N*-acetylglucosamine (GlcNAc) is a supplement which can power the hexosamine pathway and lead to the biosynthesis of uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc). UDP-GlcNAc is a substrate of the secretory pathway, leading to the assembly of complex branched *N*-glycans.

3.9. Splenocytes cocultures with MC38 cells - cytotoxic assay

As described above, splenocytes were isolated from fresh mice spleen, placed in culture in anti-CD3 coated wells and treated with KF and GlcNAc for 48h. Afterwards, splenocytes were cocultured with MC38 cells for 24h (when splenocytes were treated with GlcNAc) or 48h (when splenocytes were treated with KF). Once the coculture periods had ended, cells were incubated with anti-CD45 (primary antibody; 30 minutes on ice and protected from light; dilution 1:400), in order to differentiate immune and tumor cells, followed by a wash with FACS buffer. Then, cells were resuspended in 50 μ l of annexin binding buffer (100 mM HEPES buffer, 140 mM NaCl and 2,5 mM $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ diluted in deionized water). After this step, 50 μ l of annexin-V-APC was added and left incubating for 15 minutes (dilution 1:20; incubation at RT and protected from light) in order to bind to phosphatidylserine, which is an apoptosis marker when present on the outer cell membrane. Once this time had passed, wells were washed with annexin binding buffer and centrifuged, and then resuspended in binding buffer for collection into cytometry tubes. Just before proceeding with flow cytometry analysis, 0,5 μ l of propidium iodide (PI) - PE was added in order to bind to DNA. Samples were then acquired on FACS Canto flow cytometer and results were analyzed on FlowJo version 10 software.

3.10. Statistical analysis

Gene mRNA expression RQ values were obtained through the ΔCt method and statistical significance was determined by non-parametric one-way ANOVA (Kruskal-Wallis test). This statistical test was also used in order to determine statistical significance of MFI values from flow cytometry results. Cocultures apoptosis assay results statistical significance was determined by two-way ANOVA. *P* values equal or lower than 0,05 ($p \leq 0,05$) determined results that were statistically significant. All statistical analysis was performed using GraphPad Prism 9 software.

4. Results

4.1. Characterization of the *in situ* immuno glycoprofile along distinct colorectal cancer cascades

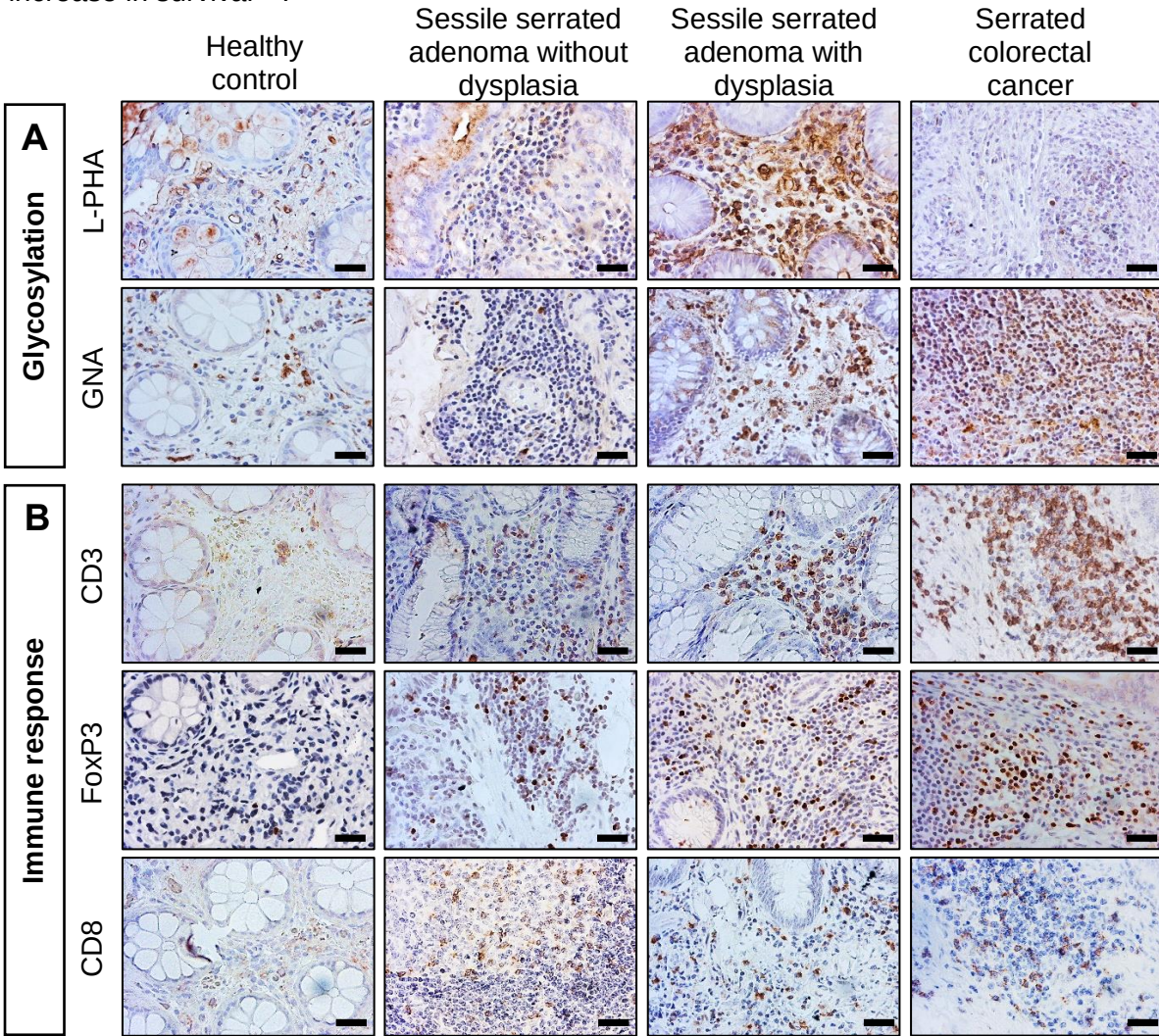
4.1.1. Serrated and Lynch syndrome CRC premalignant and cancer lesions are characterized by an increased *lamina propria* immune infiltrate that overexpresses β -1,6-*N*-glycans

Complex branched *N*-glycans are important regulators of the immune response^{63,65,71,81–84}. β -1,6-*N*-glycans are carbohydrate structures that require the intervention of *N*-acetylglucosaminyltransferase V (GnT-V), an enzyme coded by *MGAT5* gene. Our group have demonstrated these complex branched *N*-glycans to be overexpressed in colorectal cancer (CRC), unravelling a mechanism by which cancer cells escape the immune system by decreasing antigen presentation and decreasing T cell activation⁷⁸. Moreover, β -1,6-*N*-glycans are directly involved in CD4⁺ T cell activation, growth and signaling. The presence of these structures on the T cell receptor (TCR) modulates the activation threshold by decreasing receptor clustering and downstream signaling and increasing binding sites for several glycan-binding proteins (GBPs). In fact, *MGAT5* branching promotes an increase in *N*-acetylglucosamine which are the preferred binding structures for galectins, proteins that modulate T cell proliferation and apoptosis^{71,84}. Dysregulation of this mechanism, leading to a deficiency of branched *N*-glycans on the TCR, decreases T cell activation threshold and promotes a hyperactive phenotype, such as observed in ulcerative colitis⁶⁵. Nonetheless, the role of these branched *N*-glycans in the activity of T cells in CRC premalignancy and malignancy remains to be elucidated. For this purpose, we used CRC models characterized by an immense immune infiltrate.

In order to evaluate the expression profile of *N*-glycans in the immune component along the serrated and Lynch CRC pathways, we have employed histochemistry analysis on a series of formalin-fixed paraffin-embedded (FFPE) samples representing stages from both cascades. For this purpose we used two distinct lectins/GBPs: *Phaseolus vulgaris* leucoagglutinin (L-PHA), which only recognizes and binds to β -1,6-branched *N*-glycans, and *Galanthus nivalis* agglutinin (GNA), a lectin that recognizes terminal α -1,3-linked mannose. In addition, we also proceeded to characterize the specific immune profile in these lesions by employing immunohistochemistry (IHC) for T cell co-receptor CD3, cytotoxic T cell co-receptor CD8, and regulatory T cell transcription factor FoxP3.

The results of serrated lesions point to an increase of L-PHA binding (increased expression of β -1,6-branched *N*-glycans) in the immune infiltrate of sessile serrated

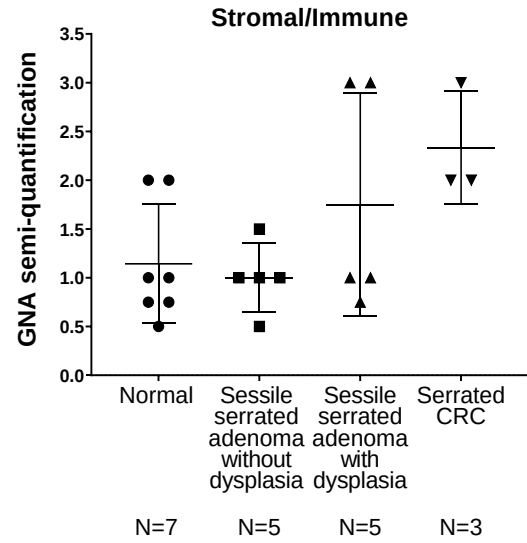
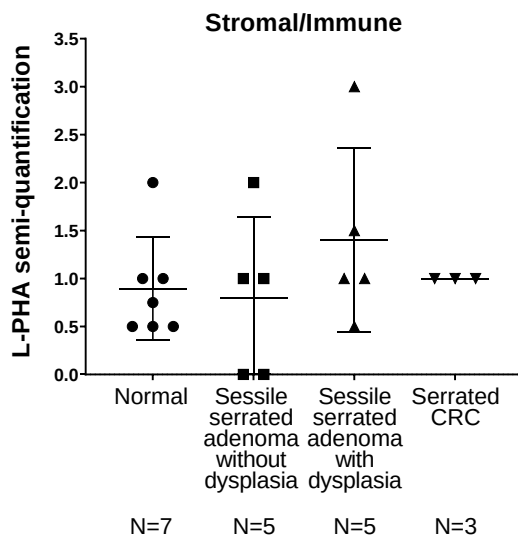
adenoma with dysplasia comparatively with sessile serrated adenomas without dysplasia and normal colonic tissue (**Figure 12A** and **Figure 12C**). Interestingly, the expression of these branched *N*-glycans decreases in cancer lesions. GNA staining (expression of high mannose *N*-glycans) in the immune infiltrate also increases along the carcinogenic pathway, peaking at the carcinoma stage (**Figure 12A** and **Figure 12C**). Furthermore, along with the progress of disease severity, an increase of infiltrate CD3⁺ T cells was observed (**Figure 12B** and **Figure 12D**). To further classify the immune phenotype of these infiltrate T cells, FoxP3 and CD8 immunohistochemistry was used to differentiate between regulatory T cells and cytotoxic T cells, respectively. Our results show an increase in regulatory T cells along the carcinogenic pathway, which is probably associated with the immunosuppressive potential of these cells and the consequent evolution to CRC (**Figure 12B** and **Figure 12D**). Interestingly, our results present a progressively steep increase of CD8⁺ tumor infiltrating lymphocytes along the cascade (**Figure 12B** and **Figure 12D**), which have been reported in the literature as predictor of good prognosis for CRC patient and an increase in survival ⁸⁵.



(figure continues in the next page)

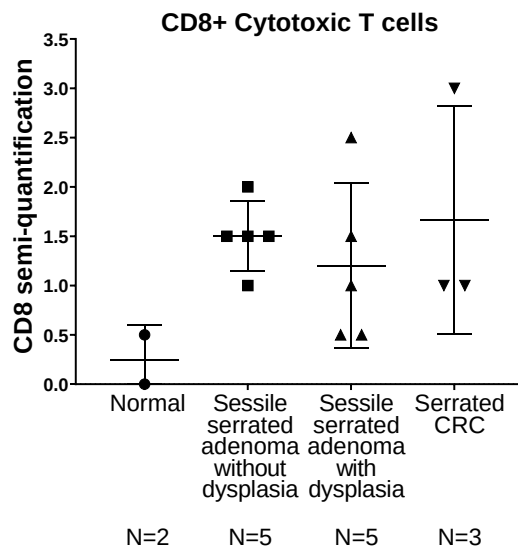
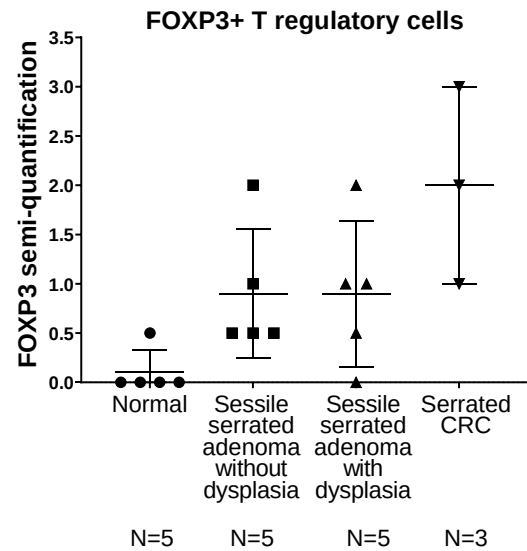
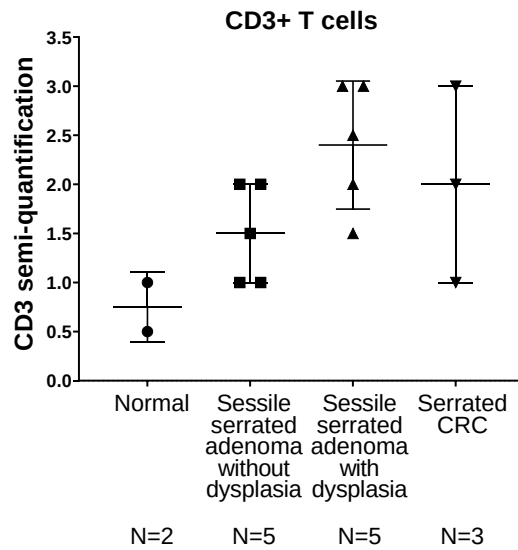
C

Glycosylation



D

Immune response



(figure legend in the next page)

Figure 12 | Evaluation of glycan and immune profile of serrated CRC pathway precursor lesions and carcinoma. Serrated pathway premalignant and carcinoma lesions were subjected to L-PHA and GNA lectin histochemistry for glycosylation profile evaluation, as well as immunohistochemistry for T cell co-receptor CD3, regulatory T cell transcription factor FoxP3 and cytotoxic T cell co-receptor CD8 to assess the immune infiltrate phenotypes. **(A)** Image panel representative of the lectin histochemistry - L-PHA and GNA (400x magnification; scale bar = 25 μ m). **(B)** Image panel representing immunohistochemistry - CD3, FoxP3 and CD8 (400x magnification; scale bar = 25 μ m). **(C)** Results from semi-quantification evaluation of L-PHA and GNA expression in the immune infiltrate of serrated lesions (performed by two independent evaluators) reveal glycome alterations throughout the pathway. **(D)** Results from semi-quantification of CD3, FoxP3 and CD8 markers in the immune infiltrate of serrated lesions (performed by two independent evaluators). The serrated cascade cohort was composed by sessile serrated adenomas without dysplasia (n=5), sessile serrated adenomas with dysplasia (n=5) and serrated CRC (n=3). Normal adjacent colonic tissue was used as control (n=2-7)

Afterwards, we aimed to validate these results by RT-qPCR on the precursor and carcinoma lesion samples and evaluate gene expression along the serrated cascade. We observed a significant increase in *MGAT5* gene expression (normalized for *18S* housekeeping) in the sessile serrated adenoma with dysplasia and serrated carcinoma stages, comparatively with normal control samples (**Figure 13A1**). This increased expression of *MGAT5* suggests an increased activity of the GnT-V enzyme and an increased surface expression of branched *N*-glycans in cells from these lesions. We also attempted to normalize the gene expression for the T lymphocyte co-receptor *CD3*, and the results obtained presented a trend increase in *MGAT5* glycoprotein expression in precursor non-dysplastic adenomas, which then seemed to decrease along the CRC pathway (**Figure 13A1**). *ST6GAL1* glycoprotein expression was also analyzed. A trend decrease in expression was observed along the cascade when normalization was performed for *18S* housekeeping (**Figure 13A2**). The same trend decrease was observed when normalization was performed for *CD3*, with a significant difference between carcinoma lesions and normal controls (**Figure 13A2**). The decreased mRNA levels of *ST6GAL1* can be associated with a decreased activity of the β -galactoside α -2,6-sialyltransferase 1 enzyme. This suggests a decreased α -2,6 sialylation (addition of terminal sialic caps) in the cells of these lesions. Furthermore, we detected a significant increase in *CD3* expression in dysplastic adenomas and cancer lesions comparatively with normal control samples (**Figure 13B**), corroborating the results from the histochemistry analysis and demonstrating an increased T cell infiltration. We also observed a tendency of increased *CD8* expression in the immune infiltrate as the cascade progressed (**Figure 13B**). Notably, we observed a trend increase in *FoxP3* gene expression in sessile serrated adenomas with and without dysplasia, which further increased in malignant lesions (**Figure 13B**).

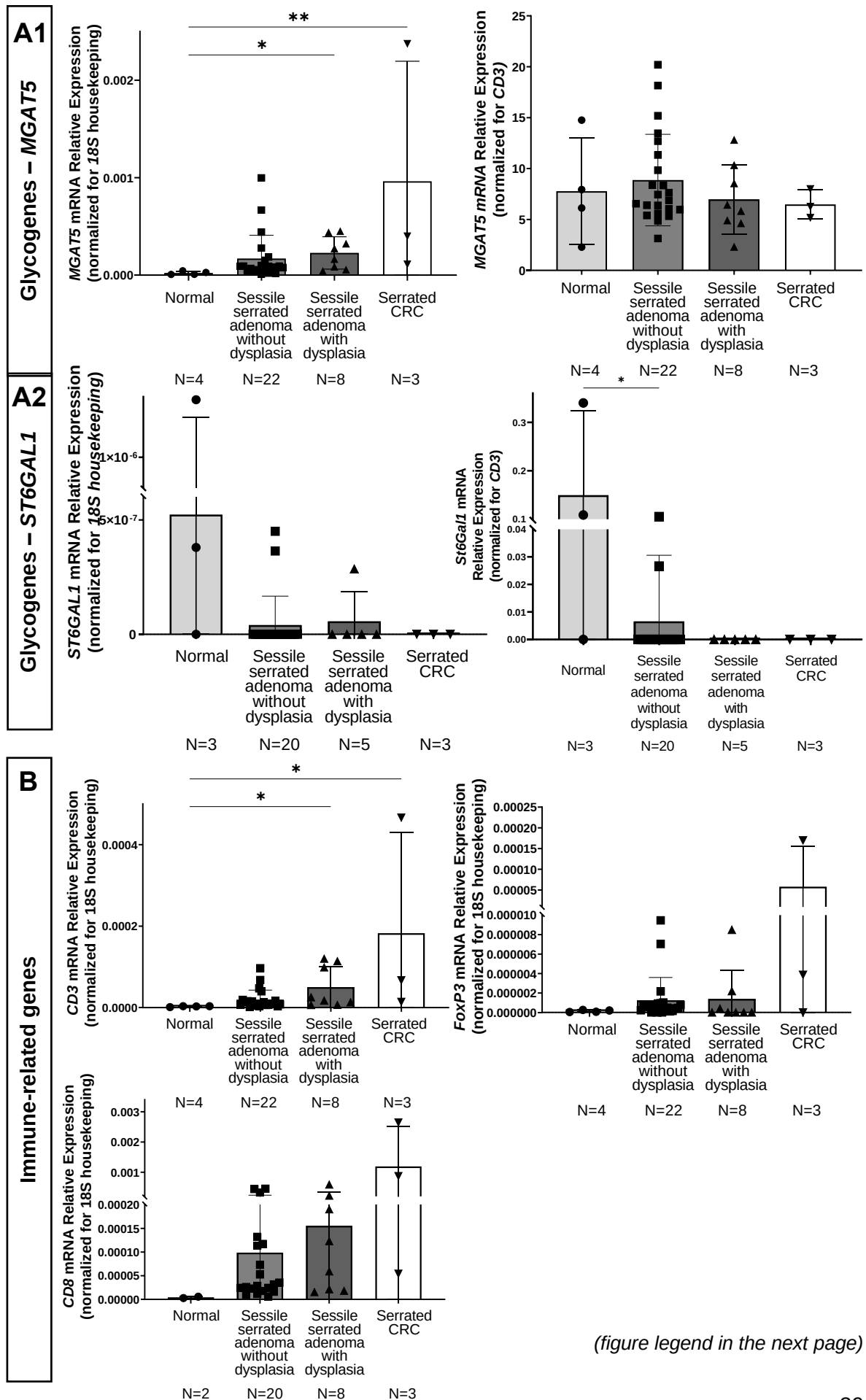
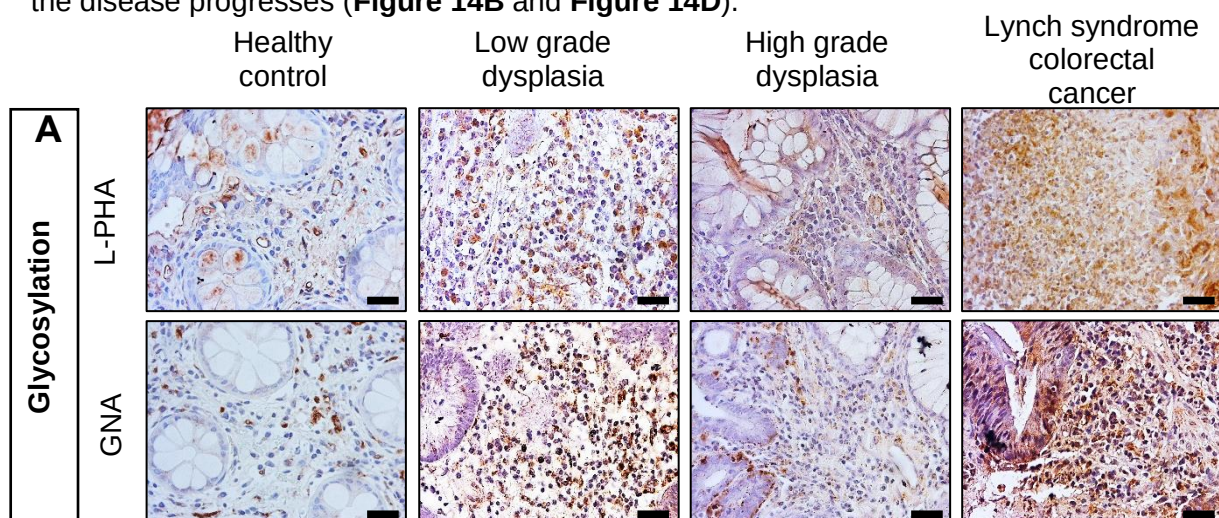
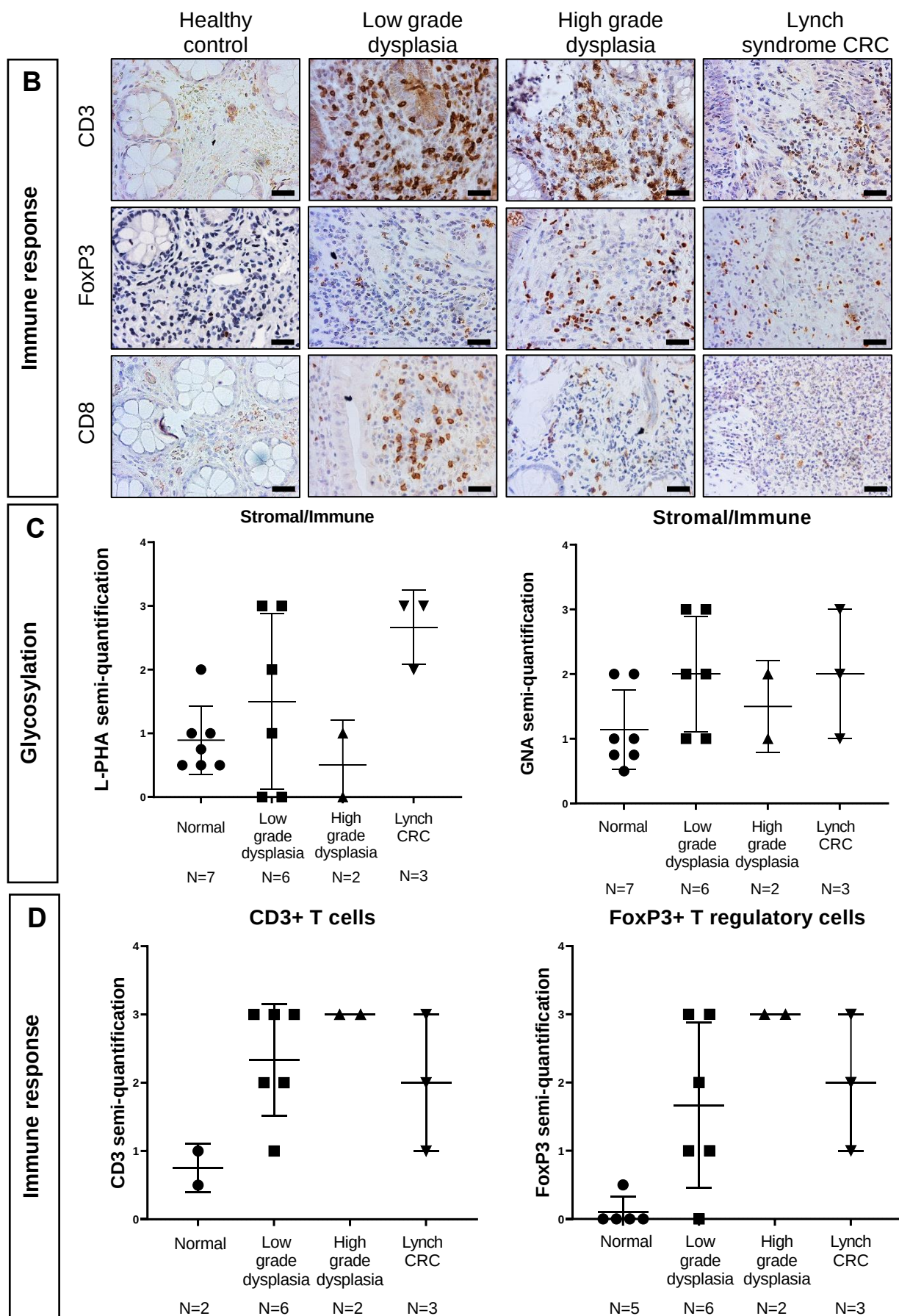


Figure 13 | Evaluation of the glyco immune profile by Real time-quantitative polymerase chain reaction (RT-qPCR) of serrated cascade precursor lesions and carcinoma. FFPE samples representative of premalignant and carcinoma lesions of the serrated CRC pathway were subjected to RNA extraction, followed by conversion into cDNA and RT-qPCR amplification. Each sample analysis was performed in duplicates and gene expression was normalized for *18S* housekeeping. Groups with $n \geq 3$ were tested for the presence of outlier values by the Robust regressing and Outlier removal (ROUT) method based on estimation of a maximum False Discovery Rate (FDR) of 1% (Q-value=1%). Statistical analysis was performed with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests: * $p \leq 0.05$; ** $p \leq 0.01$. **(A1)** *MGAT5* glycogene expression along the serrated cascade, normalized for *18S* and *CD3*, respectively **(A2)** *ST6GAL1* glycogene expression along the serrated cascade, normalized for *18S* and *CD3*, respectively. **(B)** Evaluation of immune related gene expression: *CD3* T lymphocyte common co-receptor, regulatory T lymphocyte transcription factor *FoxP3* and *CD8* cytotoxic T lymphocyte co-receptor, all normalized for *18S* housekeeping gene. Serrated cascade cohort composed by sessile serrated adenomas without dysplasia ($n=20-22$), sessile serrated adenomas with dysplasia ($n=5-8$) and serrated CRC ($n=3$). Normal colonic samples were included as controls ($n=3-4$)

We also performed L-PHA and GNA lectin histochemistry on samples representative of premalignant and carcinoma lesions from the Lynch cascade. Additionally, we analyzed the expression of *CD3*, *FoxP3* and *CD8* by immunohistochemistry to clarify the specific adaptive immune response in this pathway. Our results show that expression of branched *N*-glycans in immune infiltrate increases along the Lynch cascade (increase in L-PHA binding), with noticeable high values in the carcinoma stage (**Figure 14A** and **Figure 14C**). An increase of GNA signal (indicative of high mannose *N*-glycans) was also observed along the cascade (**Figure 14A** and **Figure 14C**). In agreement with previous results from the serrated cascade, we also perceived an increase of tumor infiltrating lymphocytes (increase in *CD3* expression), in which we could distinguish an increase in both regulatory T cells (cells expressing *FoxP3* transcription marker) and cytotoxic T cells (increased expression of *CD8* marker) (**Figure 14B** and **Figure 14D**). Remarkably, the perceived infiltration of cytotoxic T cells appears to peak at LGD lesions, with a steep decrease as the severity of the disease progresses (**Figure 14B** and **Figure 14D**).



(figure continues in the next page)



(figure continues in the next page)

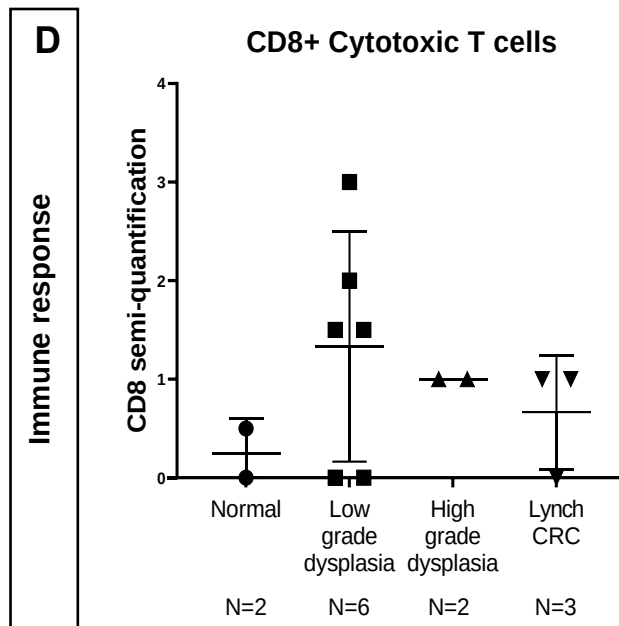


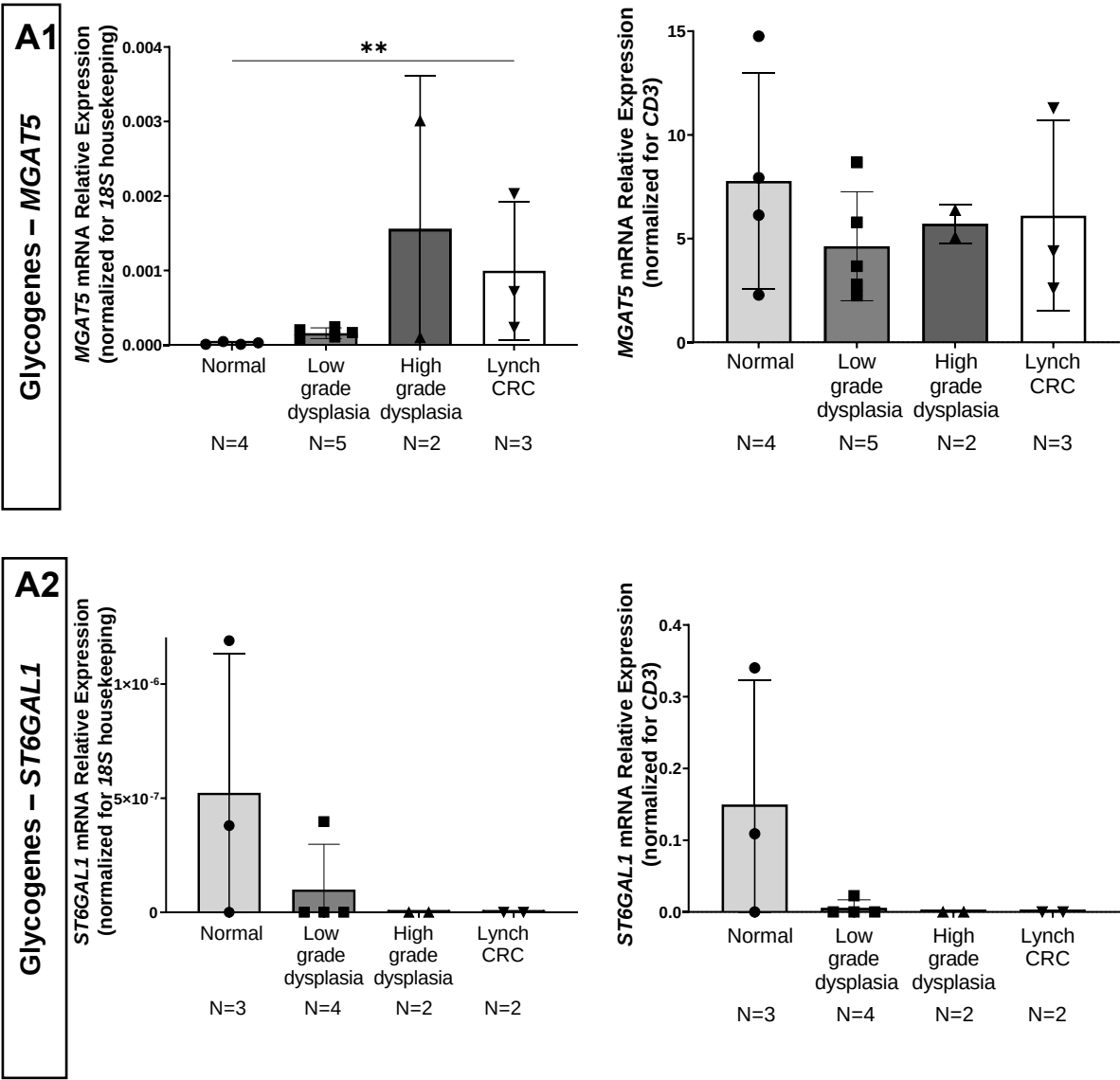
Figure 14 | Evaluation of glycan and immune profile of Lynch syndrome precursor lesions and CRC. Precursor and carcinoma lesions from Lynch syndrome patients were subjected to L-PHA and GNA lectin histochemistry for glycosylation profile evaluation, as well as immunohistochemistry for T cell co-receptor CD3, regulatory T cell transcription factor FoxP3 and cytotoxic T cell co-receptor CD8 to assess the immune infiltrate phenotypes. **(A)** Panel of images taken from representative slides (400x magnification; scale bar=25 μ m) demonstrating alterations in cell glycome - L-PHA and GNA. **(B)** Image panel representing final results from immunohistochemistry (images taken with 400x magnification; scale bar=25 μ m) - CD3, FoxP3 and CD8. **(C)** Results from semi-quantification evaluation of L-PHA and GNA expression in the immune infiltrate of serrated lesions (performed by two independent evaluators) point to alterations in the glycoprofile of immune cells **(D)** Results from semi-quantification of CD3, FoxP3 and CD8 immune markers on the infiltrate of Lynch syndrome lesions (performed by two independent evaluators). Lynch syndrome cascade cohort was composed of low grade dysplasia (LGD) lesions (n=6), high grade dysplasia (HGD) lesions (n=2) and Lynch CRC (n=3). Normal colonic samples were included as controls (n=2-7)

Moreover, RT-qPCR analysis was performed to further validate these results. We observed a trend increase in the expression of *MGAT5* glycone, with significant differences between CRC samples and normal controls (normalization for *18S* housekeeping) (**Figure 15A1**). Normalization for *CD3* immune gene was also performed, for which no significant changes were observed (**Figure 15A1**). *ST6GAL1* expression was also evaluated and was normalized for *18S* and *CD3*, and although no statistical significance was obtained, we observed that expression levels of this sialyltransferase tend to decline in Lynch pre-malignancy (**Figure 15A2**).

Focusing on the immune gene expression analysis, we observed a significant increase in *CD3* expression along the cascade, exhibiting an increased infiltration of T cells in high grade dysplasia and cancer lesions in Lynch syndrome patients (**Figure 15B**). Furthermore, we observed that this immune infiltrate increase is also correlated with higher expression of regulatory T cell markers (increase in expression of *FoxP3* transcription factor) and,

surprisingly, increased expression of *CD8* gene, suggesting an increase in infiltration of cytotoxic T cells (**Figure 15B**). This contradicted the results obtained from the immunohistochemistry evaluation (**Figure 14D**). However, it is important to note that gene expression and protein levels can be distinct and are not always correlated. Transcription is a highly regulated process. Messenger RNA (mRNA) can suffer modifications which alter its activity, such as alternative splicing or interactions with long non-coding RNA molecules. As such, gene expression does not show a perfect correlation with protein abundance. Moreover, protein translation is also controlled at the distinct phases (initiation, elongation, and post-translational modifications), and by localization and ribosome composition. Furthermore, proteins can be targeted for ubiquitination and can also be subject to interaction with other molecules, effectively influencing their abundance and activity in cells

86



(figure continues in the next page)

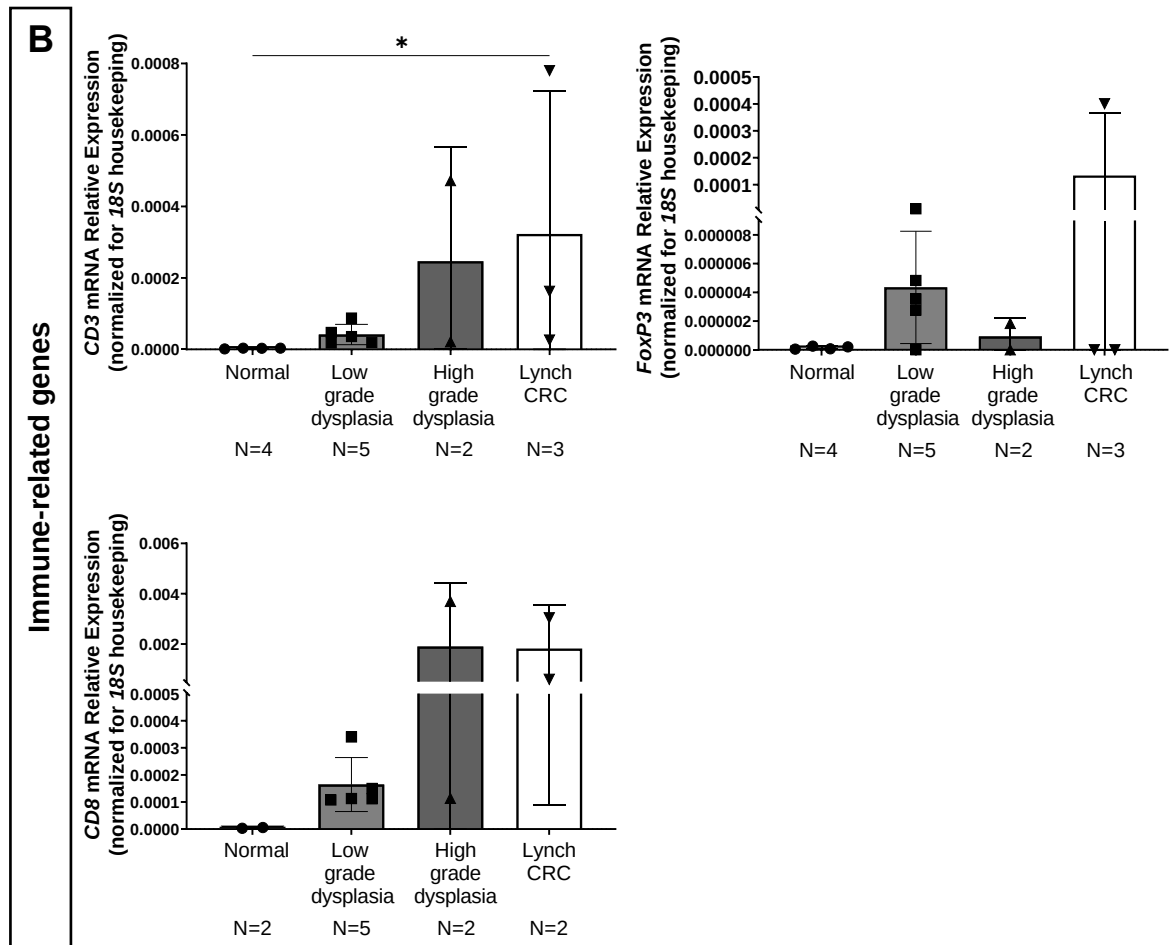


Figure 15 | Evaluation of the glyco immune profile by Real time-quantitative polymerase chain reaction (RT-qPCR) of Lynch syndrome premalignant and carcinoma lesions. RNA was extracted from FFPE samples of precursor and Lynch carcinoma, converted into cDNA and analyzed by RT-qPCR . Each sample analysis was performed in duplicates and gene expression was normalized for 18S housekeeping. Groups with $n \geq 3$ were tested for the presence of outlier values by the Robust regressing and Outlier removal (ROUT) method based on estimation of a maximum False Discovery Rate (FDR) of 1% (Q-value=1%). Statistical analysis was performed with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests: * $p \leq 0.05$; ** $p \leq 0.01$. **(A1)** *MGAT5* glycogene expression along the serrated cascade, normalized for 18S and CD3, respectively. **(A2)** *ST6GAL1* glycogene expression along the serrated cascade, normalized for 18S and CD3, respectively. **(B)** Results of immune related gene expression: CD3 T lymphocyte common co-receptor, regulatory T lymphocyte transcription factor *FoxP3* and CD8 cytotoxic T lymphocyte co-receptor, all normalized for 18S housekeeping gene. Lynch syndrome CRC cascade cohort was composed by low grade dysplasia (LGD) lesions ($n=4-5$), high grade dysplasia (HGD) lesions ($n=2$) and Lynch CRC ($n=2-3$). Normal colonic samples were included as controls ($n=2-4$).

Overall, these results suggest that the increased expression of β -1,6-*N*-glycans in the immune infiltrate of serrated and Lynch syndrome premalignant and cancer lesions is correlated with an increased infiltration of regulatory T cells. This increase in regulatory T cells is in accordance with previous findings demonstrating the important immunosuppressive action in creating an environment favorable for CRC development and

progression⁸⁷. This increased expression of branched *N*-glycans also seemed to be associated with an increased infiltration of cytotoxic T cells in the serrated cascade, but not in Lynch syndrome premalignancy and CRC. Nevertheless, the increase that we observed was puzzling and needs further clarification. As previously discussed, infiltration of these cells is correlated with a better prognosis in CRC. However, it is possible that these cells can be glycoreprogrammed to a hampered phenotype unable of mounting an effective immune response.

4.1.2. β -1,6 GlcNAc branched *N*-glycans expression appear to increase along the risk of progressing to colitis-associated colorectal cancer which is accompanied with an increase of infiltrate regulatory and cytotoxic T cells

The intense inflammation observed in inflammatory bowel disease (IBD) is also a CRC predisposing condition, with a risk of rapid evolution to cancer^{29,31,33,34}. The inflammatory response involves the participation of several immune cells, among which T cells. As previously stated, our group as already demonstrated that alterations of the T cell glycome is associated with the pathogenesis of IBD⁶⁵. Taken together, these facts indicate that colitis-associated colorectal cancer (CAC) is an attractive cascade for studying how glycans impact T cell activity in the progression from inflammation to cancer.

Regarding this cohort, we also performed lectin histochemistry to unravel glycosylation switches involved in the progression of the carcinogenic pathway. Our results suggest a trend increase in surface expression of these branched *N*-glycans along the inflammatory pathway; in other words, as the risk of colitis progresses, we observe an increase in the expression of these carbohydrate structures, which appear to peak in the intermediate/high risk colitis stages (**Figure 16A** and **Figure 16B**). Moreover, our results regarding the GNA lectin histochemistry demonstrated a trend of increased GNA binding as the inflammatory pathway progresses, suggesting an increase in high mannose *N*-glycans. However, the transition from the intermediate risk to the high risk colitis stage was defined by a decrease in GNA binding, suggesting that the progression to more advanced stages seems to be associated with a decreased expression of high-mannose *N*-glycans (**Figure 16A** and **Figure 16B**). Preliminary results from more advanced stages of the CAC cascade (which include low grade dysplasia (LGD) and CAC) seem to suggest a decreased surface expression of β -1,6 GlcNAc branched *N*-glycans in favor of high-mannose *N*-glycans (**Figure 16A** and **Figure 16B**). However, due to the small sample size, further research is required to draw significant conclusions.

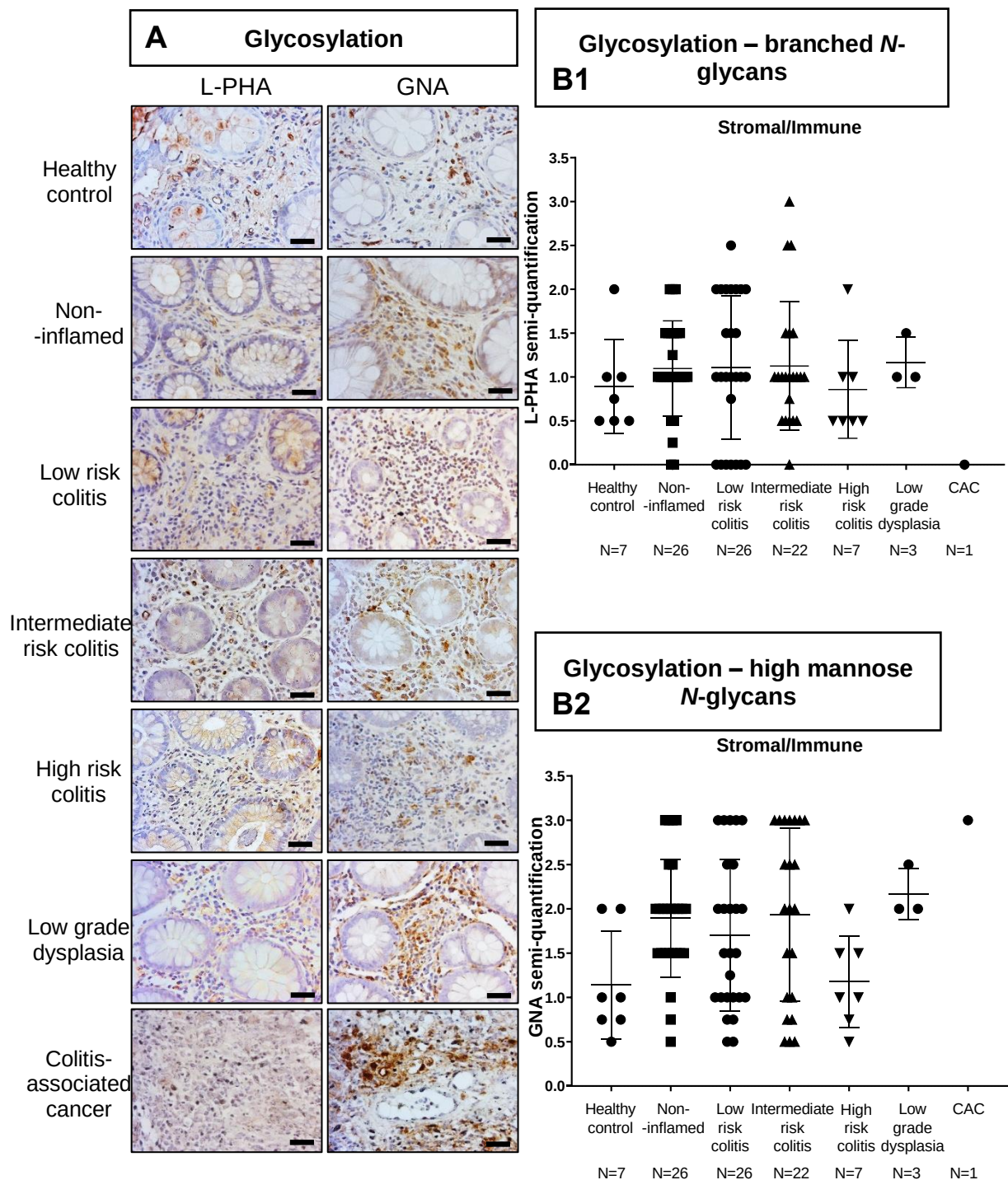


Figure 16 | Lectin histochemistry analysis of the CAC cascade immune infiltrate glycosignature. (A) Image panel representing the glycoprofile observed along the CAC cascade (400x magnification; scale bar = 25 μ m) **(B)** Stromal semi-quantification results from L-PHA **(B1)** and GNA **(B2)** lectin histochemistry performed on lesions from the CAC cascade: healthy control (n=7), non-inflamed tissue (includes adjacent tissue and tissue with macroscopic remission) (n=26), low risk colitis (n=26), intermediate risk colitis (includes polyps and Mayo 2 and 3 inflammation samples) (n=22), high risk colitis (pancolitis) (n=7), low grade dysplasia (n=3) and carcinoma (n=1). Semi-quantification was performed by two independent evaluators.

Following the lectin histochemistry analysis, RT-qPCR was also performed to assess gene expression in various stages of this cascade. Accordingly, we observed that *MGAT5* glycogene seems to be under expressed in early/low risk colitis, comparatively with normal controls (**Figure 17A**), which is in line with previous findings from our group. However, this contradicts our results from L-PHA lectin histochemistry, which do not suggest a decreased expression of β -1,6-linked branched *N*-glycans (**Figure 16A and Figure 16B**). Moreover, the expression of the gene coding for the GnT-V glycosyltransferase seems to increase from low risk to intermediate risk colitis and from intermediate to high risk colitis, suggesting an increased expression of this enzyme (**Figure 17A**). In turn, this suggests an increased expression of these glycan structures along the inflammatory pathway. Results from *MAN2A1* expression seems to corroborate this data as this glycogene is also progressively overexpressed in the inflammatory stages (**Figure 17A**). Taken together, these results seem to point to an increased expression of complex branched *N*-glycans by demonstrating that the secretory pathway is cleaving mannose from premature structures (*MAN2A1*) with further arm elongation by GnT-V (coded by *MGAT5*). Our preliminary results from LGD and CAC also suggest that these genes are under expressed comparatively with previous stages. Nevertheless, further research is required, and sample sizes need to be incremented. Furthermore, we also assessed *FUT8* and *ST6GAL1* glycogenes expression. The enzymes coded by these genes are responsible for core fucosylation and terminal sialylation, respectively. In the inflammatory pathway, we observed similar levels of expression of *FUT8* mRNA when comparing normal controls with low risk colitis samples. Afterwards, we observed a trend increase in *FUT8* expression, peaking in intermediate risk colitis (comparatively with normal controls and low and high risk colitis). This suggest an increased core fucosylation of *N*-glycan structures in intermediate stage samples. Meanwhile, *ST6GAL1* mRNA expression seems to decrease from low risk to intermediate risk colitis, and then peaks in high risk colitis, suggesting that high risk colitis is characterized by terminally sialylated *N*-glycans. In fact, *ST6GAL1* seems to be overexpressed in further advanced stages (LGD and CAC). Nevertheless, it is important to highlight once again the low sample size in these advanced stages. Furthermore, expression of *FoxP3* and *CD8* immune genes was also assessed. The results suggest an increase in regulatory T cells along the inflammatory cascade, with an apparent decrease in high risk lesions. LGD and CAC samples seem to present a trend of decreased infiltration of regulatory T cells as the disease progresses (**Figure 17B**). *CD8* gene expression revealed a peak expression in intermediate risk colitis, with a trend to decrease along the cascade and suggests a low infiltration of $CD8^+$ T cells in advanced stages (**Figure 17B**). Once again, since the number of samples in these stages was very limited, we are not able to draw significant conclusions.

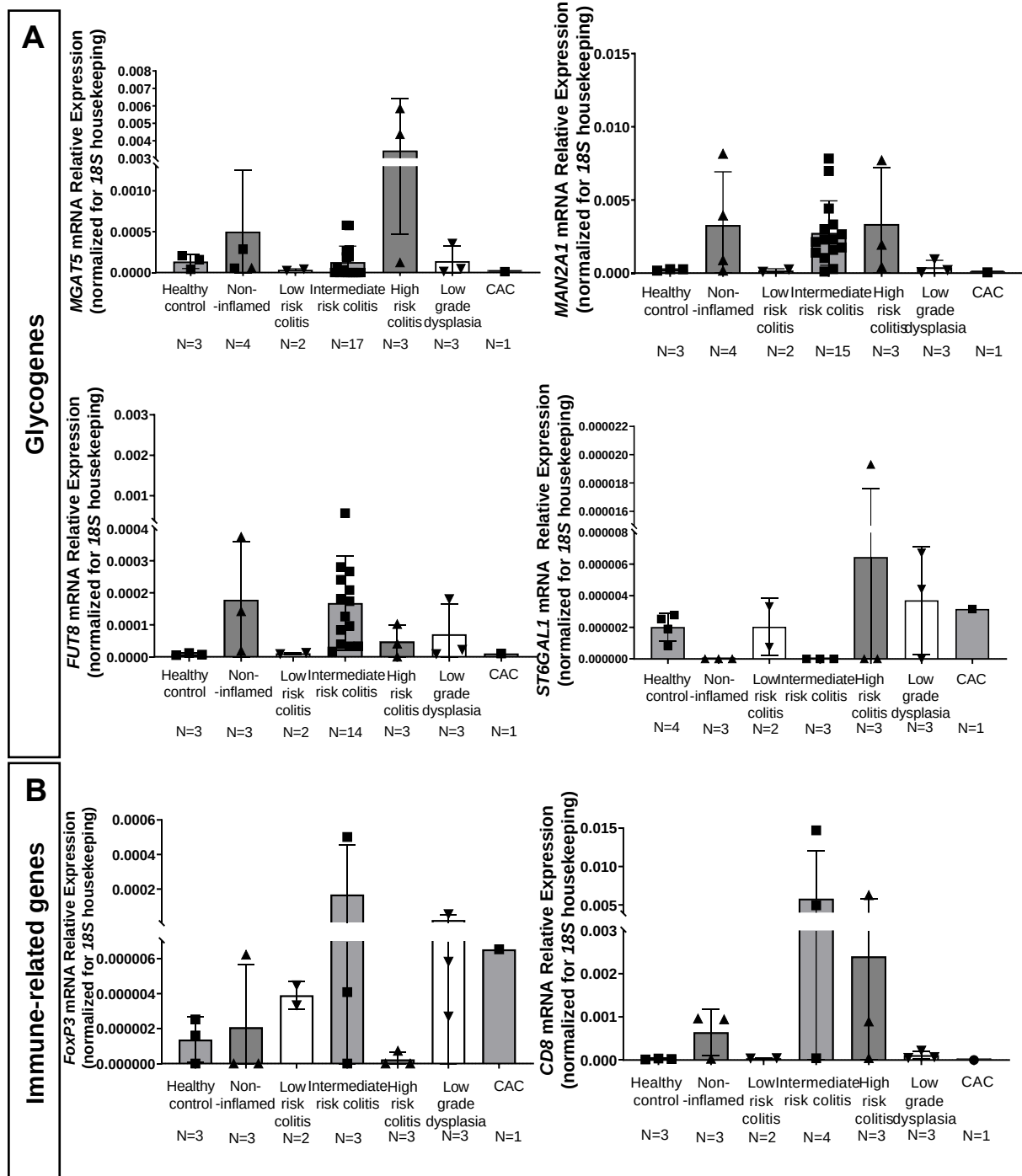


Figure 17 | Evaluation of the glyco immune profile by Real time-quantitative polymerase chain reaction (RT-qPCR) of CAC cascade lesions. Each sample analysis was performed in triplicates and gene expression was normalized for 18S housekeeping. Groups with $n \geq 3$ were tested for the presence of outlier values by the Robust regression and Outlier removal (ROUT) method based on estimation of a maximum False Discovery Rate (FDR) of 1% (Q-value=1%). Statistical analysis was performed with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests. **(A)** *MGAT5*, *MAN2A1*, *FUT8* and *ST6GAL1* glycogenes expression along CAC cascade. **(B)** Expression of immune genes coding FoxP3 transcription factor and CD8 cytotoxic T cell co-receptor. CAC cascade cohort composed by non-inflamed tissue (includes adjacent tissue and tissue with macroscopic remission) ($n=3-4$), low risk colitis ($n=2$), intermediate risk colitis (includes polyps and Mayo 2 and 3 inflammation samples) ($n=3-17$), high risk colitis (pancolitis) ($n=3$), low grade dysplasia ($n=3$) and carcinoma ($n=1$). Normal colonic samples were used as healthy controls ($n=3-4$).

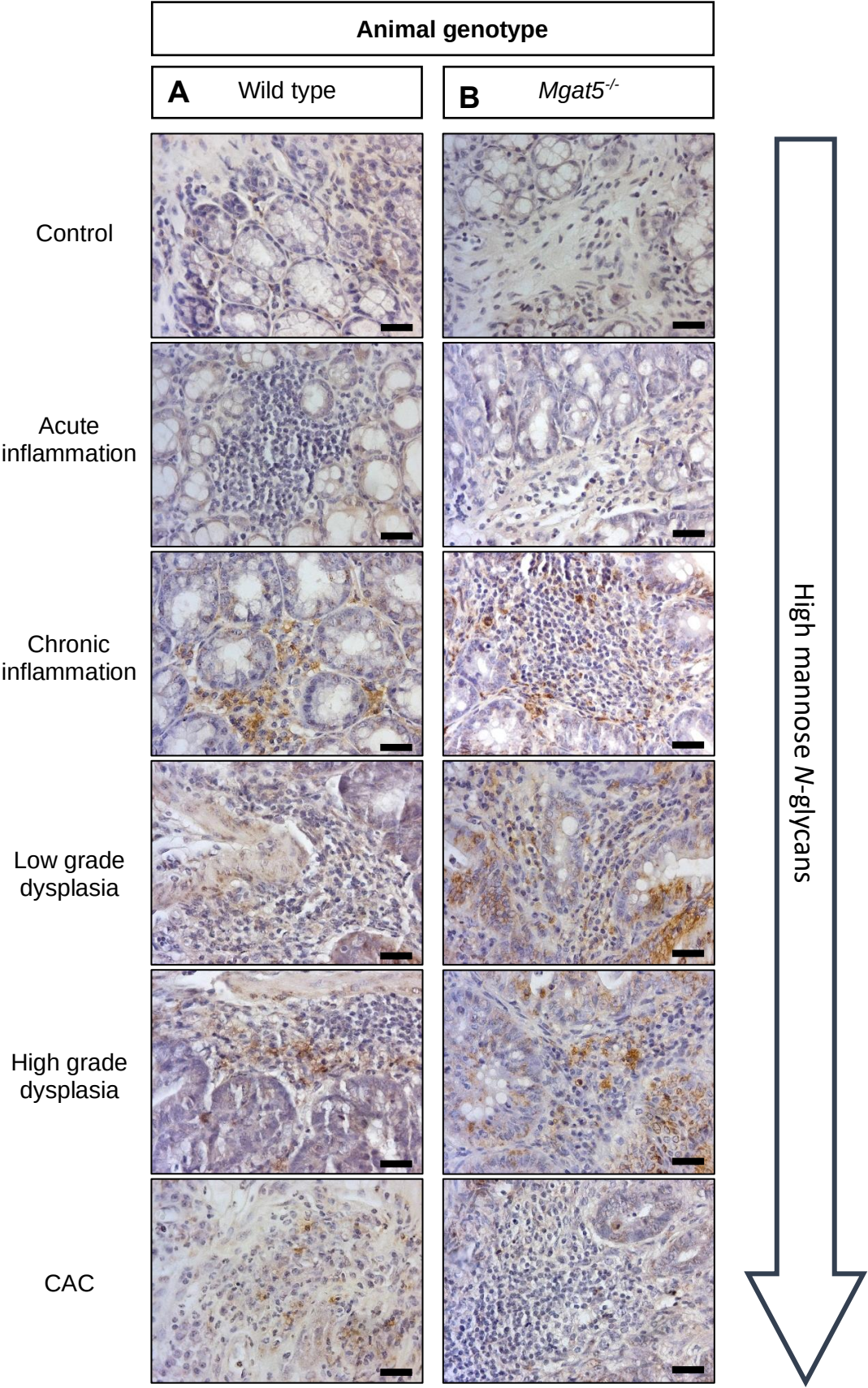
Overall, these preliminary results suggest a distinct mechanism by which glycosylation modulates the immune response and promotes progression to cancer. Immune cells infiltrating inflammatory tissue with increased risk of progression to CAC seem to be characterized by progressively increased β -1,6 GlcNAc branched *N*-glycan expression and increased terminal sialylation. This glycosignature can be associated with the increased infiltration of Tregs, explaining a mechanism that promotes immunosuppression in the transition from inflammation to cancer.

4.1.3. Mice subjected to AOM/DSS induced colitis-associated cancer exhibit a similar glycoprofile observed in human patients

In order to validate the previous results obtained from our human patient cohort, we also investigated the glycoprofile of lesions from a colitis-associated cancer (CAC)-induced mouse model. AOM/DSS is an animal model of colitis associated cancer (CAC) that remains the most useful and utilized method of mimicking the human disease⁸⁸. This model involved the administration of cycles of azoxymethane (AOM), a powerful carcinogen, and dextran sulfate sodium (DSS), a promoter of inflammation³³. In this regard, wild type (WT-*Mgat5*^{+/+}) (n=22) and *Mgat5* knockout (*Mgat5*^{-/-}, a model which does not express *N*-acetylglucosaminetransferase 5 (GnT-V) and, therefore, does not present cell surface β -1,6-*N*-glycans) (n=18) mice were subjected to induction of CAC by administration of cycles of AOM/DSS. The animal experimentation was not performed under the scope of this thesis and for this reason is not defined in the Materials and Methods section.

Colons from WT and *Mgat5*^{-/-} were collected at different timepoints and processed for histology. These samples were embedded into paraffin blocks, and lesions were characterized by an experienced pathologist. These samples were then analyzed by GNA lectin histochemistry in order to evaluate the glycoprofile of the immune infiltrate along the cascade. The results obtained from WT mice suggest an increase in GNA signal (representing high mannose *N*-glycans) in infiltrating immune cells as the disease progresses, comparatively with the control condition (**Figure 18A and Figure 18C**). These results were similar to the trend increase observed on human samples, although a decrease in high risk colitis seemed to occur (**Figure 16A and Figure 16B**). WT mice are a better representation of the disease course in a IBD patient. Nevertheless, we also performed GNA lectin histochemistry in samples from an *Mgat5*^{-/-} mice and the results show an increased expression of high mannose *N*-glycans in chronic inflammation comparatively with control samples. However, as the cascade further progresses, it appears that the

expression of these high mannose *N*-glycans decreases (although always higher than control and acute inflammation samples) (**Figure 18B and Figure 18D**).



(figure continues in the next page)

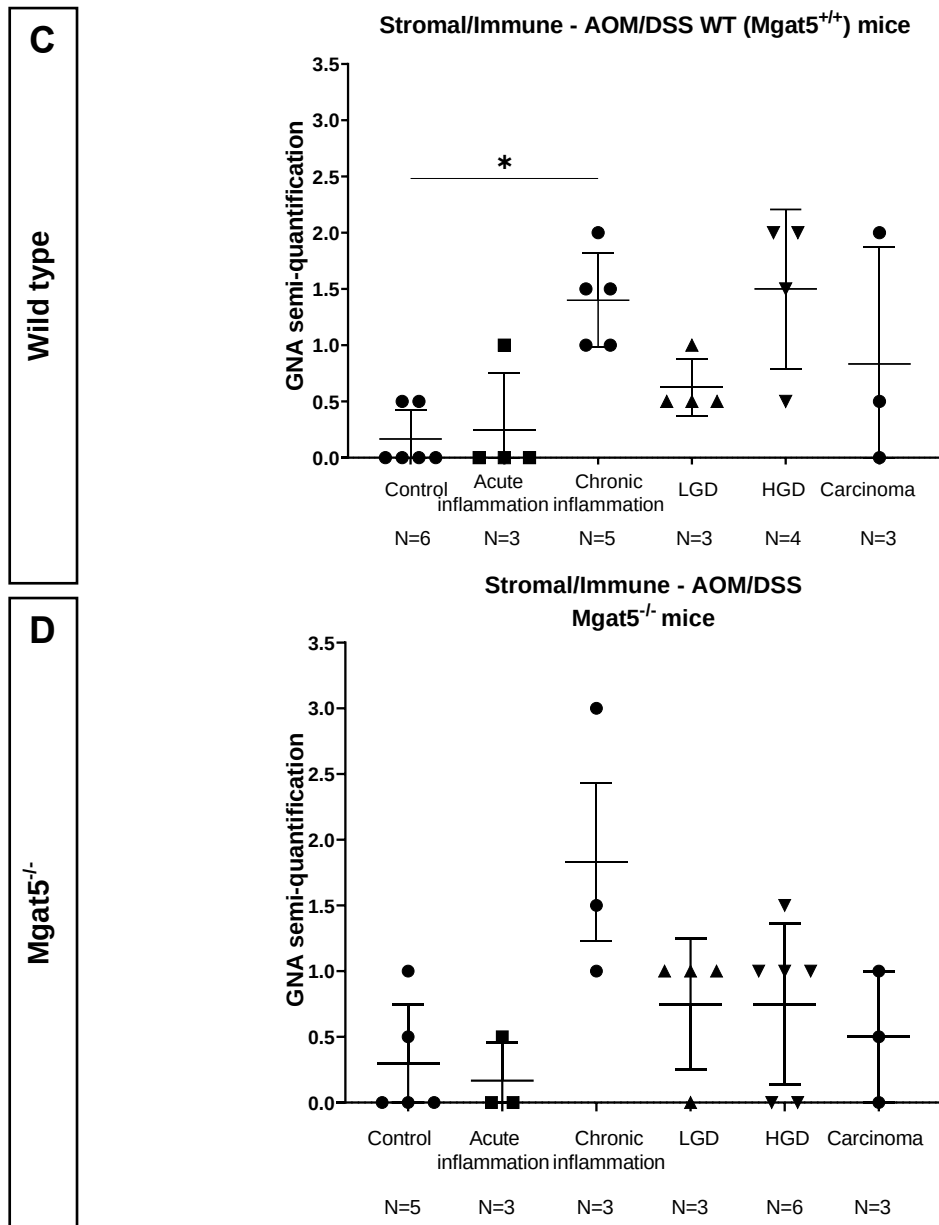


Figure 18 | Evaluation of glycoprofile by GNA lectin histochemistry performed on AOM/DSS treated WT and *Mgat5*^{-/-} mice. (A) Image panel representative of GNA signal on the immune infiltrate of a series of samples from wild type mice subjected to AOM/DSS (400x magnification; scale bar =25 μ m). **(B)** Image panel representative of GNA signal on the immune infiltrate of a series of samples from *Mgat5*^{-/-} mice subjected to AOM/DSS (400x magnification; scale bar =25 μ m). **(C)** Semi-quantification results of expression of high mannose *N*-glycans along CAC pathogenesis in Wild type mice. **(D)** Semi-quantification results of expression of high mannose *N*-glycans along CAC pathogenesis in *Mgat5*^{-/-} mice. Groups with $n \geq 3$ were analyzed for statistical significance with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests: * $p \leq 0.05$. AOM/DSS- treated WT mice cohort was composed of normal controls ($n=6$), acute inflammation ($n=3$), chronic inflammation ($n=5$), low grade dysplasia (LGD) ($n=3$), high grade dysplasia (HGD) ($n=4$) and colitis-associated colorectal cancer (CAC) ($n=3$). AOM/DSS-treated *Mgat5*^{-/-} mice cohort was composed of normal controls ($n=5$), acute inflammation ($n=3$), chronic inflammation ($n=3$), LGD ($n=3$), high grade dysplasia (HGD) ($n=6$) and CAC ($n=3$)

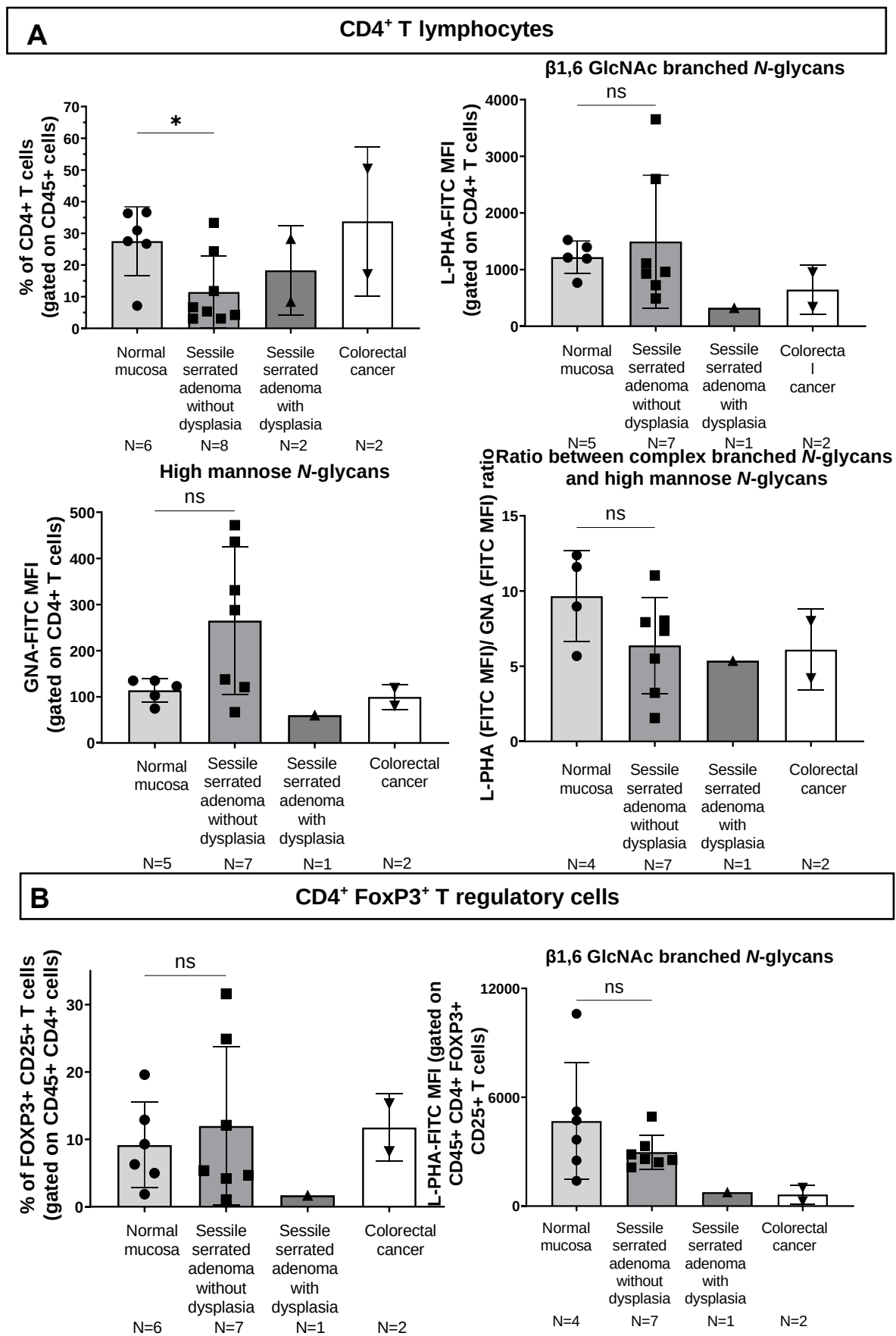
Overall, these results corroborate the data collected from the patient cohort, since the glyco profile in the inflammatory pathway seems to be similar. We observed the same increase in GNA lectin binding as the disease progressed from acute to chronic inflammation, suggesting an increase in high mannose *N*-glycans. Moreover, the use of this model allowed access to more advanced stage samples which were limited in our human cohort. The results suggest that more advanced stages of the CAC pathway seem to be characterized by decreased expression of high-mannose *N*-glycans.

4.2. *Ex vivo* analysis of T lymphocyte glycoprofile in fresh biopsies from serrated and CAC pathways colorectal cancer patients

4.2.1. Expression of complex *N*-glycans is altered in T cells along the serrated cascade – glycosylation as a modulator of the T cell-mediated immune response favoring tumorigenesis

Taken into consideration our previous results, we aimed to gain further insights into the glycoprofile changes that occur in immune cells (more specifically T lymphocytes) during the CRC development. For this purpose, fresh colon samples from serrated CRC cascade patients collected at Centro Hospitalar Universitário do Porto (CHUP): normal adjacent colonic mucosa samples (n=6), sessile serrated adenomas without dysplasia (n=8), sessile serrated adenomas with dysplasia (n=2) and colorectal cancer samples (n=2) were used. These biopsies were dissociated into single cells and subjected to FACS staining followed by flow cytometry analysis. This research involved the use of fluorescent lectins (to study the glycosylation profile) and immune cell markers to detect specific T cell populations, as formerly described.

Our results using samples from the serrated pathway demonstrate a significant decrease in CD4⁺ infiltrating T cells from normal adjacent colonic tissue to sessile serrated adenomas without dysplasia, which then increases in the following stages. Moreover, we also observed an increase of branched *N*-glycans from normal controls to sessile serrated adenomas without dysplasia, which then decreases in dysplastic and carcinoma lesions (**Figure 19A**). Surprisingly, the expression of these branched structures seems to decrease in the dysplastic and carcinoma lesions, represented as well by the L-PHA/GNA ratio (**Figure 19A**). However, it is important to note that the sample size for these two last stages is very low (n=2), which is associated with the rapid progression that occurs from sessile serrated adenomas with dysplasia to cancer ^{24,25}. To gain new insights, we will aim to increase the cohort size for these two stages. Moreover, we observed a trend increase in regulatory T cells in sessile serrated adenomas without dysplasia and CRC, comparatively with normal tissue, which was in accordance with our previous results. Tregs glycoprofile analysis revealed a tendency of decreased expression of branched *N*-glycans along the cascade (**Figure 19B**). Furthermore, the expression of cytotoxic T cells also increases along the carcinogenic pathway. In these cells we observe a decrease in GNA staining (less high mannose *N*-glycans) (**Figure 19C**).



(figure continues in the next page)

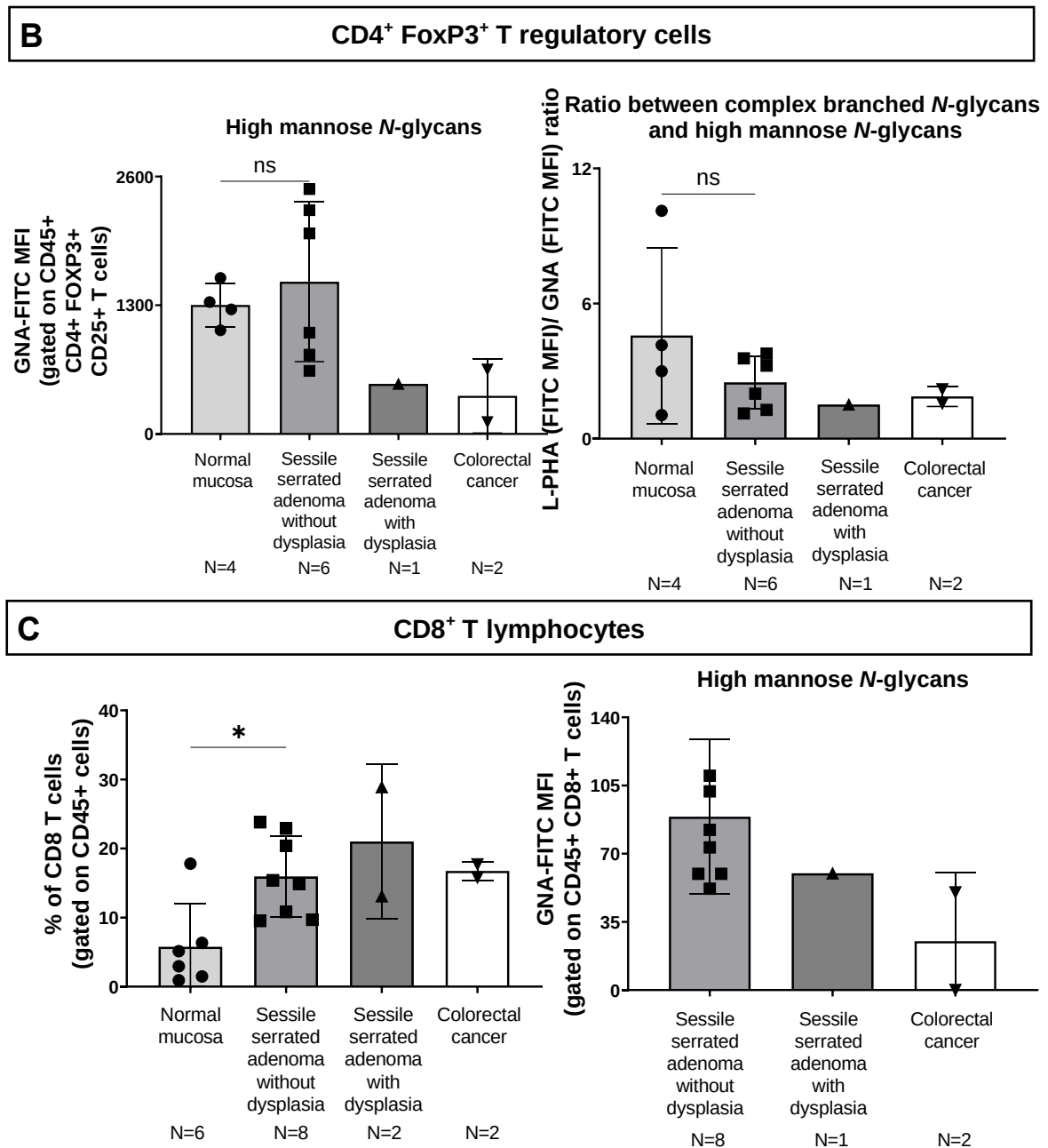


Figure 19 | Evaluation of the glycoimmune profile in premalignant serrated precursor lesions and CRC by FACS. (A) Percentage of CD45⁺ CD4⁺ T cells (gated on live single cells). MFI results from L-PHA and GNA lectin staining and L-PHA/GNA ratio. **(B)** Percentage of CD45⁺ CD4⁺ FoxP3⁺ CD25⁺ regulatory T cells (gated on live single cells). MFI results from L-PHA and GNA lectin staining and L-PHA/GNA ratio. **(C)** Percentage of CD45⁺ CD8⁺ T cells (gated on live single cells). MFI results from GNA lectin staining. Groups with $n \geq 3$ were tested for the presence of outlier values by the Robust regression and Outlier removal (ROUT) method based on estimation of a maximum False Discovery Rate (FDR) of 1% (Q-value=1%). Statistical analysis was performed with non-parametric one-way ANOVA (Kruskal-Wallis H test) and Dunn's Multiple Comparisons Tests: * $p \leq 0.05$. The cohort of serrated lesions used for FACS analysis included sessile serrated adenomas without dysplasia ($n=8$), sessile serrated adenomas with dysplasia ($n=2$) and CRC samples ($n=2$). Normal colonic samples were used as controls ($n=6$)

Overall, these results strengthen our hypothesis that branched *N*-glycans (particularly β -1,6-*N*-glycans) are positively correlated with an increased infiltration of regulatory T cells and cytotoxic T cells in premalignant lesions and CRC. These branched *N*-glycans successfully modulate T cell response, encourage recruitment or *de novo* differentiation of Tregs and lead to immunosuppression at dysplastic lesions, which is favorable for CRC progression.

4.2.2. Increased expression of branched *N*-glycans and decreased sialylation contribute for an imuno suppressive environment in the progression to dysplasia

In order to validate our previous results, we also used fresh biopsies and blood from CAC cascade patients were collected at Centro Hospitalar do Porto (CHUP) and processed for FACS stained, as previously described. Four (n=4) biopsies and two (n=2) blood samples representing inactive disease/no inflammation and dysplasia were used for flow cytometry analysis. Our results suggest less infiltrating CD4⁺ T cell in the dysplastic lesion (**Figure 20A**). Furthermore, these T cells exhibit increased branched *N*-glycans, as evidenced by increase L-PHA MFI and decreased GNA MFI (**Figure 20A**), suggesting that the evolution to premalignant lesions involves increased surface expression of branched *N*-glycans. These findings were contradictory with our previous findings, which suggested that surface expression of branched *N*-glycans decreased along the cascade. However, it is important to note that the small sample size inhibits the drawing of a significant conclusion and there is a great need to increase our cohort. No changes in terminal sialic acid expression were observed in CD4⁺ T cells (**Figure 20A**). Interestingly, blood analysis also reveals less circulating CD4⁺ T cells with a distinct glycophenotype, characterized by increased expression of branched *N*-glycans with sialic acid residues (**Figure 20B**). On the other hand, CD8⁺ T cells present similar infiltration in colonic tissue but with the same overexpression in branched *N*-glycans and decreased presence of sialic acid residues (**Figure 21A**), indicating cytotoxic T cells with increased activation threshold which can explain the lack of effectiveness in mounting an immune response to cancer cells. An increased presence of circulating cytotoxic T cells was observed in the patient with dysplasia, and these T lymphocytes shared a similar glycoprofile to the infiltrative cells (**Figure 21B**).

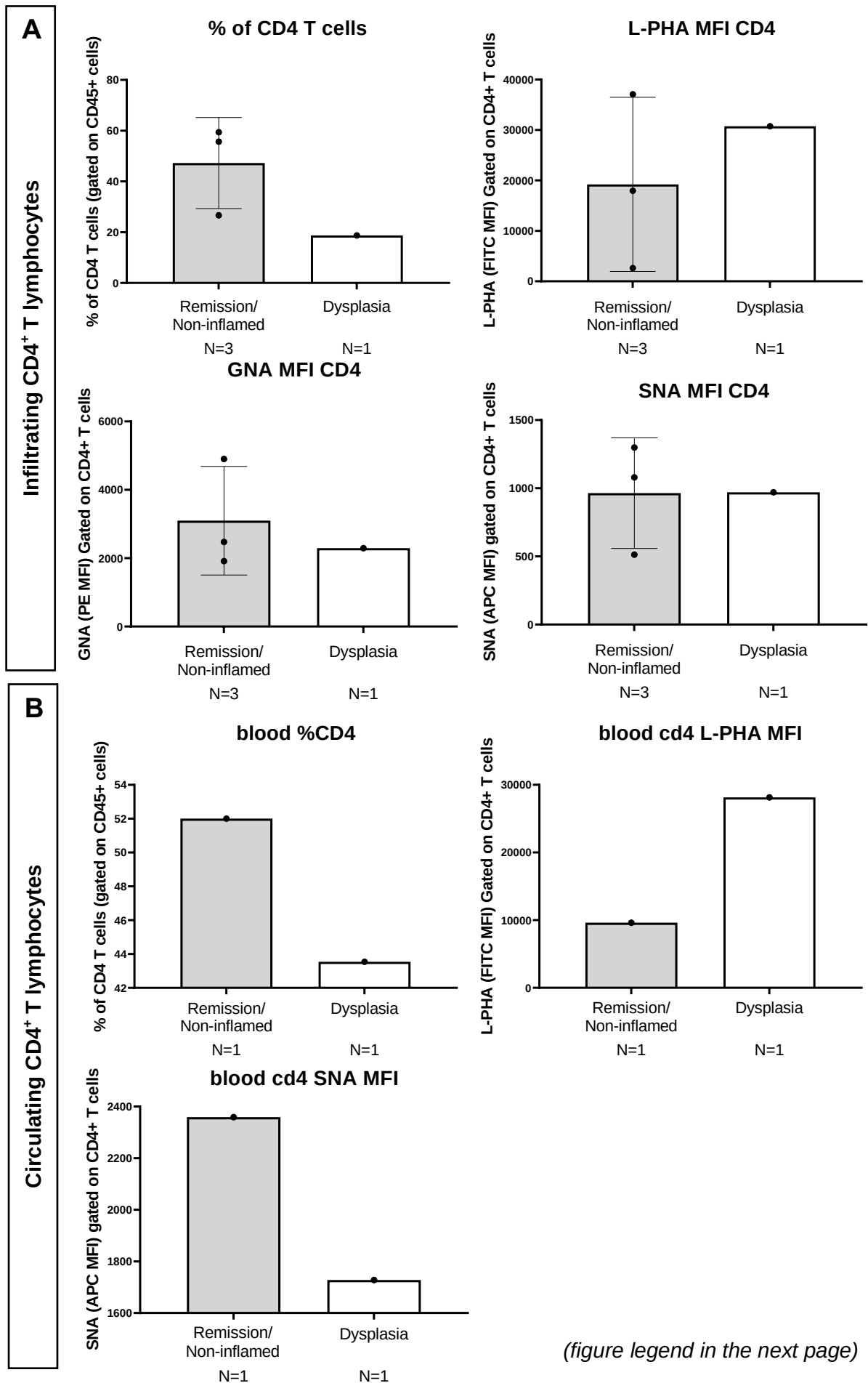
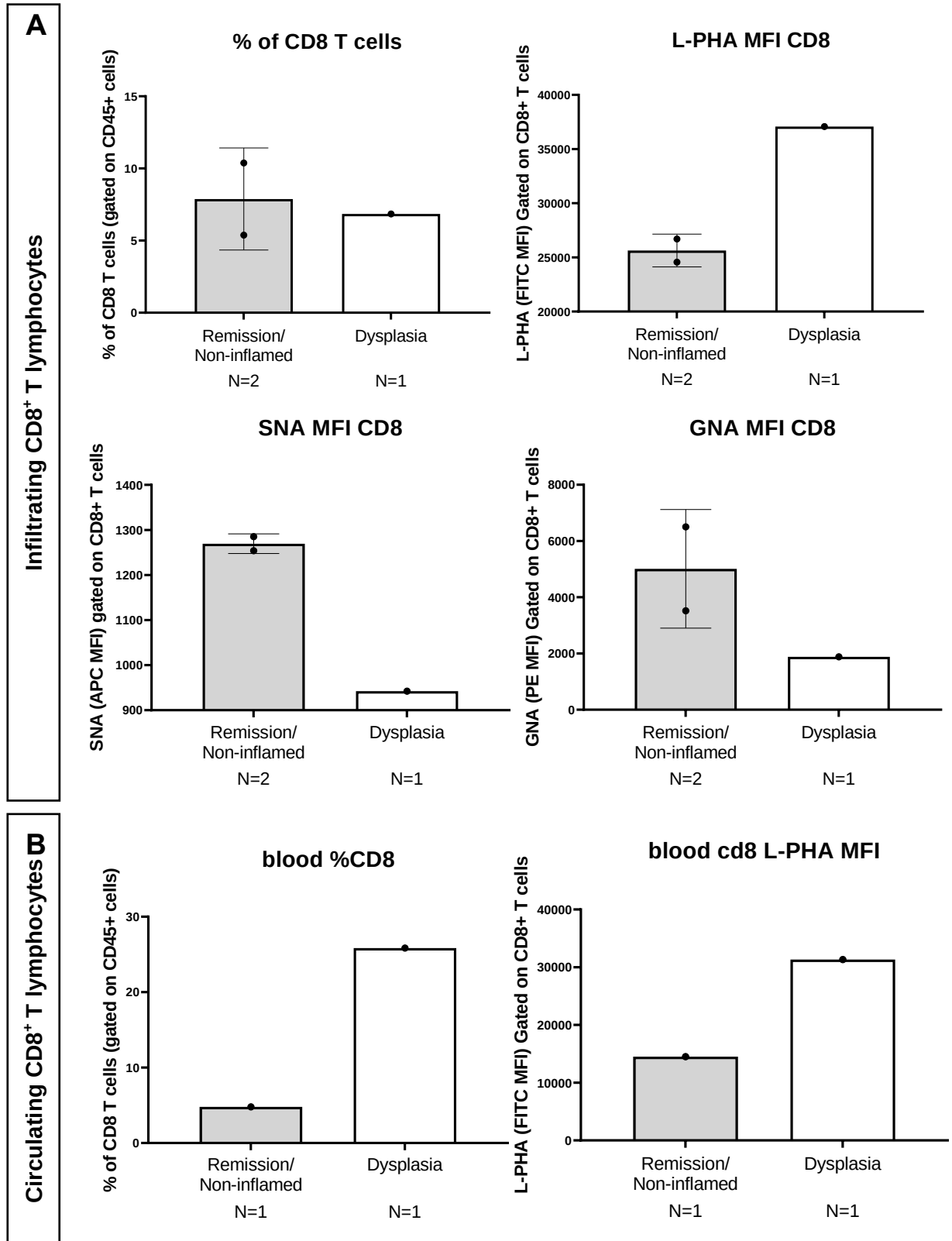


Figure 20 | CD4⁺ T cells glycoprofile in CAC fresh biopsies. (A) Percentage of infiltrating CD45⁺ CD4⁺ T cells in colon lesions. L-PHA, GNA and SNA lectin staining MFI of CD4⁺ T cells. **(B)** Percentage of circulating CD45⁺ CD4⁺ T cells. L-PHA, GNA and SNA lectin staining MFI of blood CD4⁺ T cells. No statistical analysis was performed due to the low sample size. CAC cascade stages include remission/non-inflamed adjacent tissue (n=3 for biopsies and n=1 for blood) and dysplasia lesions (n=1 for biopsies and blood)



(figure continues in the next page)

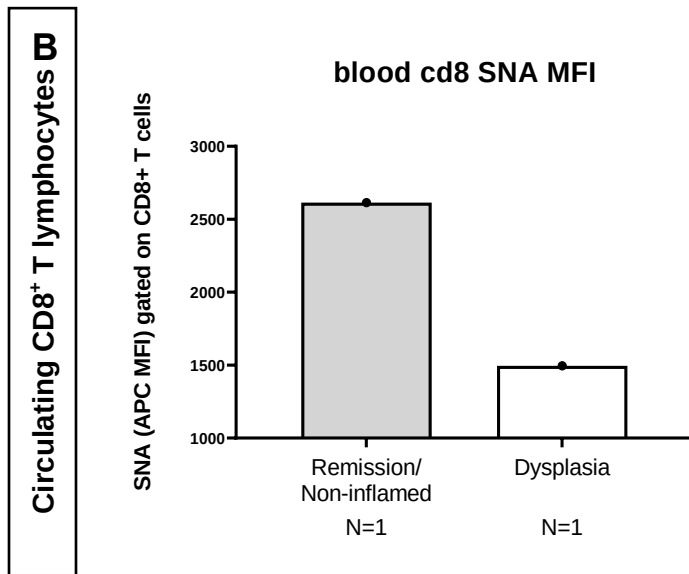


Figure 21 | CD8⁺ T cells glycoprofile in CAC fresh biopsies. (A) Percentage of infiltrating CD45⁺ CD8⁺ T cells in colon lesions. L-PHA, GNA and SNA lectin staining MFI of CD8⁺ T cells. **(B)** Percentage of circulating CD45⁺ CD8⁺ T cells. L-PHA, GNA and SNA lectin staining MFI of blood CD8⁺ cytotoxic T cells. No statistical analysis was performed due to the low sample size. CAC cascade stages include remission/non-inflamed adjacent tissue (n=3 for biopsies and n=1 for blood) and dysplasia lesions (n=1 for biopsies and blood)

Although the small sample size is an important limitation in taking significant conclusions regarding the role of *N*-glycans in T cell response in CAC, these preliminary results suggest that an aberrant expression of β -1,6-*N*-glycans on T cells occurs in premalignant stages of the CAC cascade. Since this sample size is very small, there is a need to increase the number of samples and disease stages. Moreover, sialic acid residues capping glycan structures are inhibitors of interactions with several GBPs, such as galectins. Galectins can promote T cell apoptosis and suppression ⁸⁹. In this regard, the decreased expression in sialic acids that we observe might in turn explain how immune cells are remodeled in order to promote a tolerant and propitious environment for CAC cells to thrive.

4.3. T lymphocyte glycoreprogramming: metabolic supplementation inducing overexpression of β -1,6-*N*-glycans inhibits effective immune response against cancer cells

As previously discussed, β -1,6-*N*-glycans deficiency plays an important role in several autoimmune disorders by lowering TCR signaling threshold and inducing a hyperactive T cell response^{65,83,84}. In order to better understand the role of these branched *N*-glycans in the response against CRC cells, we designed an *in vitro* approach based on the evaluation of effective targeting of CRC cells by glycoreprogrammed T cells. For this purpose, we used MC38 cells (a murine adenocarcinoma cell line) as targets for the immune response. Splenocytes were extracted from a wild type (WT) mouse spleen and subjected to metabolic supplementation with either kifunensine or *N*-acetylglucosamine (GlcNAc). Kifunensine is an inhibitor of α -mannosidase I enzyme and blocks the *N*-glycosylation pathway, promoting the expression of high mannose structures in favor of complex branched *N*-glycans⁹⁰. On the other hand, GlcNAc supplementation can supply metabolites for the hexosamine pathway, which is responsible for UDP-GlcNAc biosynthesis, an essentially building block for assembly of branched *N*-glycans in the Golgi apparatus⁹¹. As demonstrated by Ana Dias *et al.*, GlcNAc supplementation is an effective therapeutic approach for ulcerative colitis (UC) patients, by decreasing T cell inflammatory activity and decreasing disease severity⁷⁴. **Figure 11** (Material and Methods section) is a graphical representation of the effects of these two supplements in cell glycome.

Initially, we tested at least two concentrations of supplement (except for kifunensine, whose concentration of 20 μ M was already optimized in our laboratory) and distinct timepoints (24 hours, 48 hours and 48 hours + 24 hours without supplementation) to assess T cell metabolic reprogramming. Afterwards, these cells were stained for flow cytometry analysis in order to evaluate their glycoprofile. Our results showed that kifunensine inhibitory action is already observable after 24 hours of incubation with the immune cells, as demonstrated by the decrease in L-PHA MFI, further decreasing after 48 hours and even after 24 hours without supplementation (**Figure 22A**). In accordance, GNA and ConA (a lectin recognizing all mannose residues in a glycan structure) MFI progressively increased, validating the inhibition of branched *N*-glycans in favor of high mannose *N*-glycans (**Figure 22B and Figure 22C**). Interestingly, when we evaluated the results of L-PHA MFI in splenocytes treated with 100 mM GlcNAc, we observed a slight decrease in L-PHA MFI (decrease in β -1,6-*N*-glycans) after 24 hours of treatment, which was unexpected, as GlcNAc should increase production of UDP-GlcNAc and increase expression of these branched *N*-glycans (**Figure 22D**). This suggests that supplementation with GlcNAc might require more time until producing detectable changes. After 48 hours of treatment, levels of

L-PHA MFI had increased and, after 48 hours of supplementation and further 24 hours without GlcNAc in the medium, L-PHA MFI had increased beyond the levels from the baseline control (**Figure 22D**), demonstrating an effective action by GlcNAc in stimulation expression of β -1,6-*N*-glycans. This suggests a delayed action by this supplement in inducing an observable response, which can be attributed to the fact that GlcNAc has to enter the cell, power the hexosamine pathway in order to produce UDP-GlcNAc and only after this can it be used as a substrate by the Golgi *N*-acetylglucosaminetranferases.

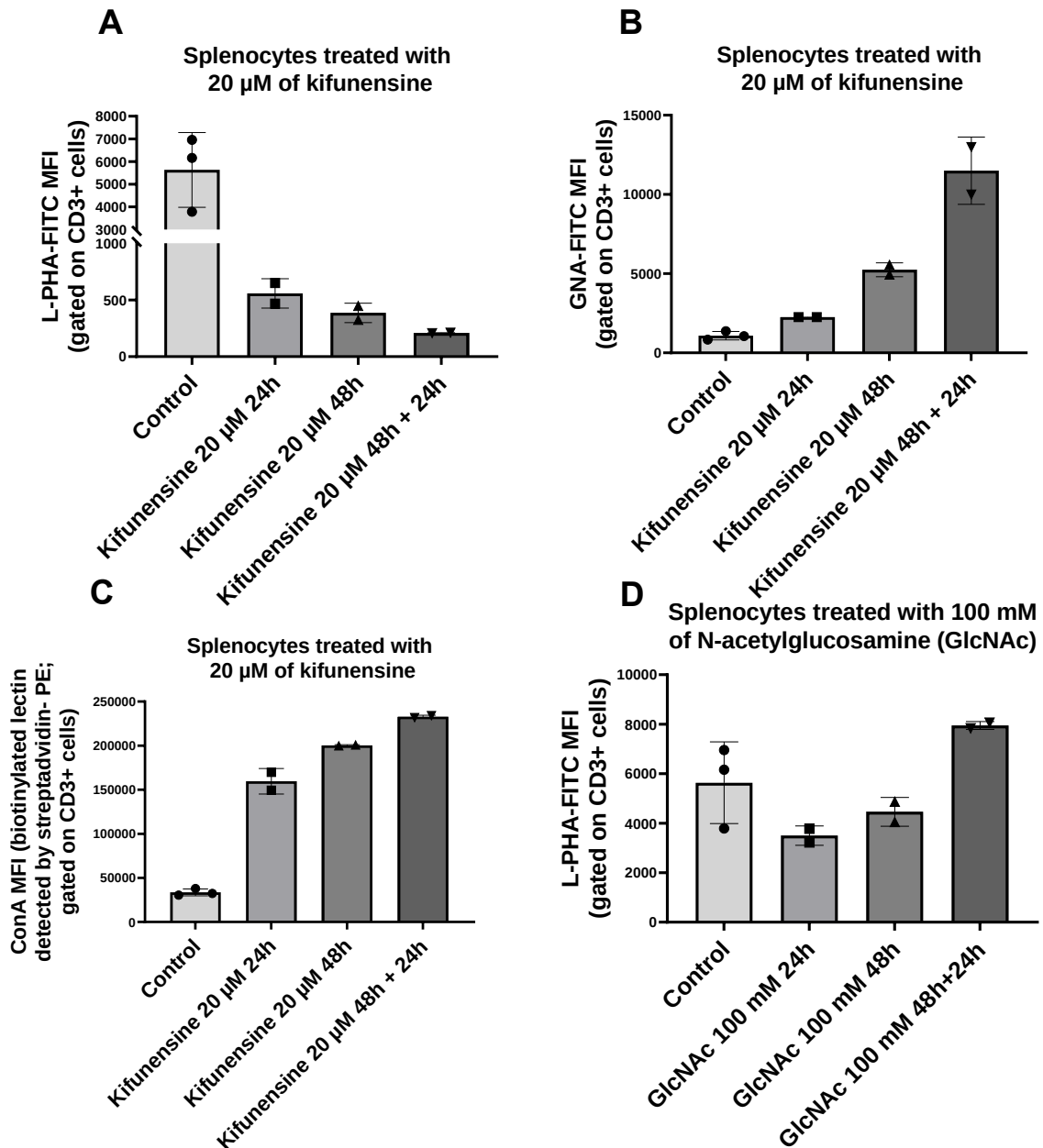


Figure 22 | Metabolic supplementation and T cell reprogramming optimization experiment. (A) L-PHA staining (FITC MFI), **(B)** GNA staining (FITC MFI) and **(C)** ConA staining (biotinylated lectin detected by streptavidin- PE MFI) of T cells cultured with 20 μ M of Kifunensine for 24, 48 and 48 hours + 24 hours of recovery compared with untreated splenocytes. **(D)** L-PHA staining (FITC MFI) of T cells cultured in medium supplemented with 100 mM of GlcNAc for 24, 48 and 48 hours + 24 hours of recovery compared with untreated control. No statistical analysis was performed due to the low sample size ($n \leq 3$).

Splenocytes were also treated with 200 mM of GlcNAc. However, this concentration presented high toxicity, culminating in the death of all cells (analyzed by flow cytometry – data not shown). Regarding this data, we decided to continue our *in vitro* experiment using 20 μ M of kifunensine and 100 mM of GlcNAc treatments over 48 hours + 24 hours for recovery, as this timepoint proved to be ideal for both treatments and would be applicable for the experiment design (splenocytes would be treated for 48 hours and then place in cocultures with the cancer cell line for 24 hours without supplementation). However, our data regarding 24h cocultures revealed no observable differences between kifunensine-treated splenocytes and non-treated splenocytes (data not shown), and as such we decide to evaluate the efficacy in 48h cocultures. Afterwards, the viability of the MC38 cells was measured by flow cytometry (annexin V/propidium iodide (PI) assay). With this method, we were able to verify the efficacy of the immune response in relation to the glycoprofile induced by metabolic supplementation. Notably we observed a trend decrease in live MC38 cells cocultured with KF treated splenocytes (18.47%) comparatively to the same population percentage when coculture with untreated splenocytes (27.13%) (**Figure 23A** and **Figure 23B**), suggesting that in fact KF is impacting the interaction between immune and cancer cells. We also observed a trend increase in apoptotic and dead cancer cells (percentages in **Image 23B**).

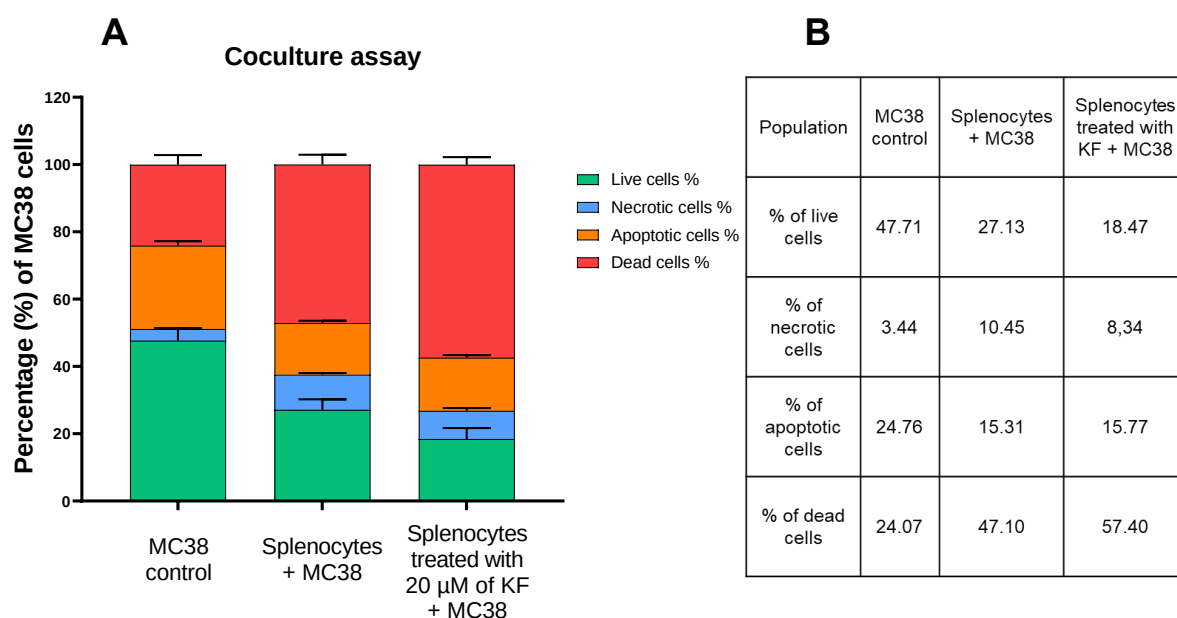


Figure 23 | MC38 murine adenocarcinoma cells viability after coculture with splenocytes untreated or treated with kifunensine. Cocultured cells were stained for FACS analysis. Cancer cells were distinguished from immune cells by CD45 staining and evaluated for cell death by annexin-V/PI method in order to assess the efficacy of splenocyte immune response after glycoreprogramming. No statistical analysis was performed. **(A)** Graphic plot of apoptosis assay percentages. **(B)** Mean percentages of each population distinguished by annexin-V/PI. This experiment was conducted with MC38 control cells (n=3), splenocytes + MC38 cocultures (n=2) and KF-treated splenocytes + MC38 cocultures (n=3).

Importantly, GlcNAc treated splenocytes were less effective in promoting MC38 cells death, as observed by a significant increase in live cancer cells when comparing samples cocultured with treated T cells (% of live cells =34.53) and untreated T cells (% of live cells=25.33) (**Figure 24A** and **Figure 24B**). These results suggest that GlcNAc-induced β -1,6-*N*-glycan branching is promoting an inactive T cell phenotype, as also shown by the decrease in necrotic, dead, and apoptotic MC38 cancer cells (percentages in **Figure 24B**).

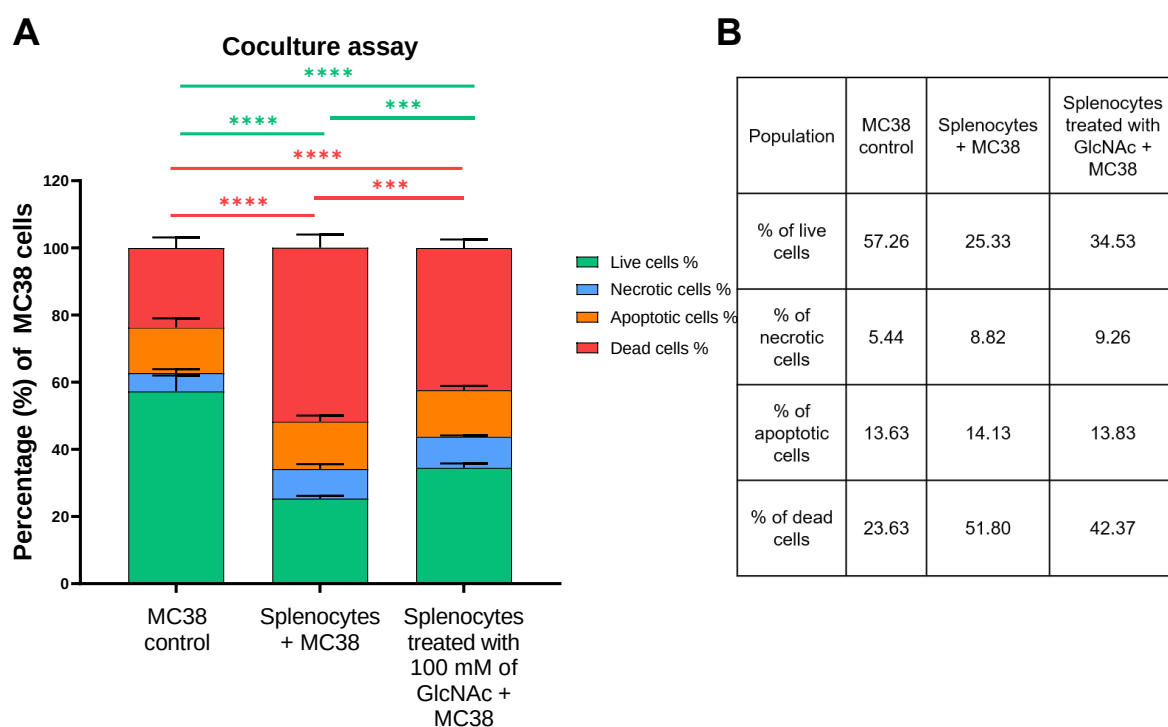


Figure 24 | MC38 murine adenocarcinoma cells viability after coculture with splenocytes untreated or treated with N-acetylglucosamine. Cocultured cells were stained for FACS analysis. Cancer cells were distinguished from immune cells by CD45 staining and evaluated for cell death by annexin-V/PI method in order to assess the efficacy of T cell response after glycoreprogramming. **(A)** Graphic plot of apoptosis assay percentages. Statistical analysis was performed by an ordinary two-way ANOVA with multiple comparisons: *** $p \leq 0.001$; **** $p \leq 0.0001$. **(B)** Mean percentages of each population distinguished by annexin-V/PI. This experiment was conducted with MC38 control cells (n=3), splenocytes + MC38 cocultures (n=3) and GlcNAc-treated splenocytes + MC38 cocultures (n=3).

Overall, our results have demonstrated that β -1,6-*N*-glycans are important mediators of cell-cell interaction. Immune cells, namely T cells require this contact with target cells in order to survey for threats and to initiate an immune response. We have therefore validated that these complex branched *N*-glycans are associated with a decreased T cell activity and decreased cancer targeting, effectively allowing the proliferation of the malignant cells (as demonstrated by the increase in viable cancer cells when T cells are glycoreprogrammed to overexpress β -1,6-*N*-glycans). Moreover, removal of these complex *N*-glycans is associated with a surveillant T cell phenotype with increased effectiveness in killing cancer

cells (as demonstrated by the decrease in viable cancer cells when they are cocultured with T cells glycoreprogrammed to display high mannose *N*-glycans). This glyco switch on the T cells (overexpressing β -1,6-*N*-glycans) may be occurring in the tumor microenvironment and could present an explanation for CRC cell immune escape.

5. Discussion

Colorectal cancer (CRC) is a major public health issue, as its incidence and mortality are increasing at alarming rates ^{1,10,11,15–17}. This highly heterogeneous disease is attributed to multiple risk factors and can progress via multiple carcinogenic pathways, namely the adenoma-carcinoma sequence, the serrated pathway and the inflammatory pathway ¹¹. Although novel diagnostic methods, allied with better implemented screening programs, have been useful to detect early stage disease, a significant percentage of patients are diagnosed with advanced stage cancer, compromising their long term survival ^{13,14}. Moreover, there is a great need to research novel biomarkers and therapeutic targets in order to decrease the burden of CRC.

Inflammatory bowel disease (IBD) is a highly debilitating disease characterized by intense inflammation of the colon and whose incidence has been steadily increasing worldwide ^{29,30}. IBD patients also present an increased risk of developing colitis-associated cancer (CAC), a malignancy arising in the colonic inflammatory environment with distinct characteristics from sporadic CRC ^{29,32,34,35}. Since IBD is a chronic disease continuously affecting increasing numbers of patients, preparations must be done in order to anticipate the possible increase in CAC cases. Glycosylation is an important mechanism in cellular homeostasis and has severe implications in disease ^{26,71}. The immune response is heavily reliant on glycosylation and, as such, dysregulation can lead to autoimmune diseases and cancer ^{60,62,63,75–77,81}. Our group has already demonstrated that ulcerative colitis (UC) patients exhibit a particular T cell glycoprofile (with decreased *MGAT5*-mediated branched *N*-glycans), which influences T cell receptor (TCR) signaling and promotes a hyperactive inflammatory response ⁶⁵. More recently, new findings from our group have elucidated novel glycome changes in CRC cells, which promote immune system evasion ⁷⁸. In this regard, glycosylation has presented huge potential in being an innovative research target for elucidating autoimmune and cancer mechanisms of pathogenesis. Moreover, deciphering changes in cell glycome can unravel new therapeutic approaches to relieve the burden of these diseases.

The present study addresses how differential T cell glycosylation pattern is involved in the progression from normal mucosa to premalignancy/CRC. For this purpose, we extensively studied the T cell glycome in different subtypes of CRC and in its precursor lesions, including serrated and Lynch syndrome CRC and CAC, characterized by an intense immune infiltrate at tumor sites. With this, we intend to suggest the glycoreprogramming of T cells as an innovative immunotherapy for CRC prevention and treatment.

Our results from serrated and Lynch syndrome have demonstrated that branched *N*-glycans (more specifically, *MGAT5*-mediated β -1,6-linked branched *N*-glycans) are overexpressed in T cells infiltrating in premalignant lesions. Moreover, the presence of these branched *N*-glycans on the cell surface is suggested to influence the progression to CRC by promoting a less effective T cell-mediated response. Moreover, this seems to be correlated with an increase in expression of Foxp3 and CD8, as demonstrated by immunohistochemistry and validated by RT-qPCR, demonstrating an increase of T regulatory cells (Tregs) and cytotoxic CD8⁺ T cells. While an increase of Tregs is in line with the current understanding of how these immunosuppressive cells are associated with a progression to cancer, an increase of infiltrating cytotoxic cells has been associated with a better prognosis for CRC patients^{85,92}. Although the results mentioned seem contradictory, it is known that Th2 lymphocytes and CD4⁺ CD25⁺ FOXP3⁺ Tregs are able to decrease CD8⁺ T cells TCR signaling and antigen sensitivity through an immunosuppressive pathway involving IL-10-mediated upregulation of *MGAT5* and consequent increased expression of branched *N*-glycans⁹³. In this sense, CD8⁺ recruitment to the tumor site might occur in premalignant and CRC lesions, but modifications to its glycome by the abovementioned populations (which are in fact Th phenotypes induced by increased expression of branched *N*-glycans on CD4⁺ T cells) can interfere with effectiveness in cancer targeting. Finally, this increased CD8⁺ immune infiltrate can be explained by galectin-1-mediated reprogramming promoting differentiation into the less researched CD8⁺ Tregs phenotype, as this has recently been demonstrated to also play a role in the immunosuppression in CRC⁹⁴. We also validated the results of serrated lesions (premalignant and cancer) using fresh biopsies by flow cytometry, and our preliminary results pointed to a decrease in branched *N*-glycans in infiltrating CD4⁺ T cells and CD4⁺ CD25⁺ FOXP3⁺ Tregs along the cascade. However, it is important to note that the number of samples needs to be incremented, mainly in sessile serrated adenomas with dysplasia. These lesions are difficult to detect in the screening colonoscopies, mainly due to their flat profile and enormous potential for rapid progression to CRC^{24,25}. In accordance with the data obtained from histochemistry and RT-qPCR, we also observed an increase in the percentages of infiltrating Tregs and cytotoxic T cells. We propose that the dysregulation in the T cell glycome is associated with the promotion of immunosuppressive networks, which create a thriving environment for CRC cells to grow, proliferate and disseminate. β -1,6-linked branched *N*-glycans are further elongated by poly-*N*-acetylglucosamine (poly-LacNAc) residues, which are binding sites for galectin molecules. Galectins are proteins with high affinity for *N*-glycans rich in β -galactose⁹⁵. These glycan-binding proteins (GBPs) are important regulators of several cellular mechanisms: cell migration, adhesion, proliferation, apoptosis, and intracellular glycoprotein transport. Galectins are implicated in autoimmunity, inflammation and,

importantly, cancer ^{71,89,95}, by modulating immunosuppressive mechanisms ⁷¹, such as suppression of effector T cell response, T cell exhaustion and promotion of differentiation into T regulatory cells ⁸⁹. Increase of β -1,6-linked branched *N*-glycans leads to increased presence of motifs that bind these GBPs. Galectin-3 is a GBP that forms a lattice responsible for modulation of TCR signaling and control of receptor clustering, reducing antigen sensitivity ^{83,84}. Galectin-1 is also involved in this immunosuppressive network and has been demonstrated to be produced and secreted by cancer cells and modulate the immune response by triggering differentiation towards a Th2-mediated cytokine profile, favoring cancer progression ^{82,87,96}. Moreover, galectin-1 sensitizes T cells for induced apoptosis, once again demonstrating how these molecules can exert T cell immunosuppression ^{87,96}. Furthermore, Tregs and Th2 (but not Th1 and Th17) cells are protected from galectin-1 action by increased expression of terminal α -2,6-sialic acid caps, leading to an increased polarization of this tumor promoting T cell phenotype ^{87,96,97}. In accordance, our *ST6GAL1* glycogene expression analysis showed a decreased on its expression in serrated and Lynch premalignant/CRC lesions, when compared with normal adjacent colonic tissue. Overall, the results we obtained, and the current literature seem to suggest a possible involvement of the galectin-glycan interaction in the progression from normal mucosa to premalignancy/CRC by promoting T cell suppression. In addition, β -1,6-linked branched *N*-glycans are known to favor membrane localization of immune inhibitory checkpoints, such as PD-1 and CTLA-4, and are required for interaction with the respective ligands, demonstrating once again that complex branched *N*-glycans participate in the creation of immunosuppressive networks ^{72,98–100}. Increased *N*-glycan branching and consequent increased surface retention of these checkpoint inhibitors on Tregs also increase their suppressive potential ¹⁰¹. Therefore, this data validates our findings and suggest a strong correlation between the increase of β -1,6-linked branched *N*-glycans and the increased presence of Tregs in premalignant lesions and cancer. Overall, our results are indicative of a high importance of β -1,6-linked branched *N*-glycans in the modulation of T cell activity and implication in CRC. Impressively, the immuno glycoprofile changes are already observed in premalignant lesions, suggesting that this reprogramming occurs before progression to carcinoma, and indicate a great potential as a target for preventive therapy. However, there is a great need to increase our sample size, primarily in fresh biopsies of sessile serrated adenomas with dysplasia and carcinoma in order to obtain reliable information regarding the glycome alterations occurring in these lesions.

On the other hand, our results regarding CAC suggest that infiltrating T cells are characterized by increased branched *N*-glycans surface expression along the inflammatory pathway, as demonstrated by L-PHA lectin histochemistry and *MGAT5* glycogene RT-

qPCR. Moreover, we observed low expression of high mannose *N*-glycans (as demonstrated by decreased binding of GNA lectin and increased *MAN2A1* glycogene expression) along this pathway. Taking into consideration that decreased *N*-glycan branching in T cells is associated with an hyperactive immune response^{65,83,84}, and that our RT-qPCR results suggest a decreased expression of *MGAT5* in low risk colitis, comparatively with normal controls, we validate that these lesions present an immune infiltrate with a pro-inflammatory glyco profile (as previously demonstrated by our group). Inflammation is often correlated with tumor initiation through production of reactive oxygen species (ROS), mutation of genomic and mitochondrial DNA, increased cell instability, activation of oncogenes and inactivation of tumor suppressor genes^{29,32,33}. However, surface expression of β -1,6-linked branched *N*-glycans seems to progressively increase along the inflammatory pathway (from low risk to intermediate risk to high risk colitis), suggesting that the risk of progression to malignancy might be associated with the presence of this particular type of *N*-glycans. Our evaluation of GNA lectin binding on samples from AOM/DSS colitis-induced mice model suggest that the expression of high-mannose structures actually increased along the inflammatory pathway. We observed a similar trend when transitioning from normal mucosa to low risk colitis and from low risk to intermediate risk colitis. However, our data regarding the human cohort suggested that a decrease occurs in high risk colitis, which was not observable on chronic inflammation samples from AOM/DSS-induced mice. We also performed flow cytometry analysis of colonic and blood samples from a remission patient and a patient exhibiting dysplastic lesions. Although the sample size is very low, the profile seems to be in accordance with our hypothesis: there seems to be a decrease infiltration of CD4 and a similar infiltration of CD8 T cells, which are characterized by increased expression of β -1,6-linked branched *N*-glycans and decreased terminal sialylation. Also regarding the T cell profile, we observed a trend increase in expression of *FOXP3* along the inflammatory stages, peaking in the intermediate stage while being apparently absent in high risk colitis. Taking in consideration all of our results and the state of the art, we hypothesize that the pro-inflammatory and hyperactive T cell phenotype apparently promoted by a decreased expression of *MGAT5*-mediated β -1,6-linked branched *N*-glycans observed in early colitis might switch progressively along the severity of the inflammation. As this increase occurs, binding sites for galectin-3 and galectin-1 may also increment. The aberrant expression of binding motifs for galectin-3 is a key aspect in the clustering of TCR and increased T cell activation threshold^{83,84} and can be responsible for decreased T cell activity and promotion of Tregs differentiation and recruitment in the progression of CAC. Moreover, galectin-1 binding can also increase when terminal branch sialylation decreases. As we observed in our RT-qPCR results, *ST6GAL1* expression seems to increase along the cascade. This would be in with our hypothesis

regarding the increased infiltration of Tregs, since Tregs these cells are usually protected from galectin-1 mediated apoptosis by terminal sialylation^{87,96,97}. Therefore, we suggest that several glyco switches occur along this pathway and are responsible for the progression to CAC. As our group has already demonstrated, the decreased expression of *MGAT5* and consequent decrease in surface expression of β -1,6-linked branched *N*-glycans on T cells is the factor driving the initiation of inflammation in the intestinal mucosa^{65,71,102}. This represents the first step in the cascade leading to CAC. Afterwards, there seems to occur an increase in the levels of these branched *N*-glycans in the surface of T cells along the inflammatory pathway. This may represent a glyco switch, which might be responsible for the transition to advanced stages. Moreover, we observed once again that aberrant expression of β -1,6-linked branched *N*-glycans seems to be associated with increased infiltration of Tregs and increased immunosuppression, and this mechanism is a possible explanation for the progression to CAC.

In order to unravel the impact of β -1,6-linked branched *N*-glycans in the interaction between T cells and cancer cells, we performed cocultures involving MC38 murine adenocarcinoma cell like with kifunensine (KF) or *N*-acetylglucosamine (GlcNAc) treated T cells. Interestingly, we observed that KF treated T cells (which exhibit high mannose *N*-glycans) were slightly more efficient in destroying the cancer cells. This proves that β -1,6-linked branched *N*-glycans are important mediators of T cell activity, influencing in the surveillance and destruction of cancer cells. Moreover, these results suggest an interesting preventive therapy target for patients at risk of developing CRC. *Ex vivo* glycoreprogramming of T cells to induce a highly efficient phenotype to resolve these premalignant lesions is an attractive model to prevent progression to CRC without implications of immune rejection and toxicity for these patients. Interestingly, the potential of glycoreprogramming in cancer has been explored before. Swainsonine is a potent inhibitor of the Golgi apparatus α -mannosidase II enzyme leading to the synthesis of hybrid *N*-glycans (oligosaccharides with one high-mannose arm and one complex-type arm)^{103–106}. As previously pointed, as glycans have immunomodulatory properties and play a key role in cancer. Cancer cells are also characterized by a dysregulated glycome that confer advantageous properties, such as metastatic potential and survival¹⁰⁶. Altering these cells glycome was proposed was an interesting target to reduce cancer invasiveness. Phases I and IB of the first clinical trials revealed a successful shrinkage of solid tumors and improvement of disease-related symptoms. The common side effects detected on the patient cohort were increase serum retinol, increased serum amylase, mild liver dysfunction and edema. The main problem arising from the treatment with swainsonine was toxicity to the liver, particularly in patients liver metastases. In fact, hepatotoxicity was the dose-

limiting toxicity. Other complications involved acute respiratory distress syndrome in a patient with hepatic dysfunction, which lead to death^{103,105}. No further phases or trials with swainsonine were conducted.

In conclusion, we obtained important results that provide information regarding the modulation of the immune response (namely T cells, which were the scope of this thesis) by *N*-glycans in serrated, Lynch syndrome and CAC cascades. Glycosylation is at the interface between immunity-cancer, with implication in how the disease will progress. Importantly, β -1,6-linked branched *N*-glycans play a role in T cell response, influencing the activity of this immune cells and promoting immunosuppression, being overexpressed in serrated and Lynch syndrome premalignancy/CRC and in the initial stages of the inflammatory pathway of the CAC cascade. Furthermore, we showed that β -1,6-linked branched *N*-glycans are key modulators of the interaction between T cells and cancer cells. *In vitro* glycoreprogramming T cells to express surface less complex/high mannose *N*-glycans appear to contribute to cancer suppression, unravelling a novel target for directed/personalized immunotherapies for CRC. Our results indicate kifunensine as a potential treatment option to promote T cell action against CRC. As no clinical trials have been conducted to evaluate the effect of kifunensine *in vivo*, further research on the impact of this molecule and glycosylation on the interface between immunity and cancer are required.

6. Conclusion and future perspectives

With the present work we consolidate the importance of glycosylation in the modulation of the T cell-mediated immune response. Particularly, β -1,6-GlcNAc branched *N*-glycans were demonstrated to be key regulators of T cell interaction with cancer cells, decreasing the effectiveness of immune surveillance and reducing targeting of cancer. This hypothesis was validated using human samples from serrated and Lynch syndrome. Serrated and Lynch syndrome cascades seem to involve a glycoreprogramming of infiltrating T cells, leading to increased expression of *MGAT5*-mediated *N*-glycan branching. Moreover, this glyco switch also seems to be associated with an increased presence of regulatory and, surprisingly, cytotoxic T cells. Further studies need to be performed to validate these promising results, including increasing sample size for histochemistry, RT-qPCR and fresh biopsies. Galectin-1 histochemistry is also an interesting experiment that would explore the presence of binding sites and validate some of the hypothesis that we proposed in the discussion. Furthermore, our results could also be validated by the use of biological material from AOM-induced CRC mice models with MMR-loss (mimicking Lynch syndrome cascade) and CIMP-high/BRAF mutations (representative of serrated cascade). Interestingly, the CAC cascade results also suggest that β -1,6-GlcNAc branched *N*-glycans expression increases along the inflammatory pathway as the risk of progression of colitis increases. This glycoprofile also seems to be associated with recruitment of Tregs to the lesion site, promoting the growth and proliferation of epithelial cells, which have been modified due to prolonged exposure to inflammatory molecules. As such, there seems to be an intricate relation between T cell glycome and immunosuppression and this seems to be a key factor in CAC. Due to the limited sample availability, we were not able to evaluate the T cell glycome and phenotypes along more advanced stages of the disease (low grade dysplasia, high grade dysplasia and carcinoma). This represents a limitation of this work, which needs to be resolved by increasing the numbers of biopsies and blood samples. In the scope of this work, we attempted to gather information regarding these advanced stages by using samples from AOM/DSS-induced mice models (mimics the CAC cascade). The results obtained support the data obtained from the human cohort, suggesting that the expression of high-mannose *N*-glycans increases in the inflammatory cascade, but decreased in more advanced stages. Since this is a well-established model, colons extracted from mice at different timepoints with distinct disease stages could be used for further analysis by L-PHA lectin and immunohistochemistry, RT-qPCR and FACS analysis to evaluate the specific immuno glycoprofile in premalignant/CAC lesions. Additionally, another interesting approach could involve the adoptive transfer of T cells from wild type (WT) and *Mgat5*^{-/-}

mice to AOM/DSS-treated immunodeficient mice. With this model, we could evaluate the evolution to CAC based on the glyco profile of the transferred cells. This would include the monitorization of the disease course in the animals and could unravel how the expression of β -1,6-GlcNAc branched *N*-glycans on T cell surface is associated with the pathogenesis of CAC.

Considering our *in vitro* results regarding T cell glycoreprogramming, it would also be interesting to move to CRC and CAC cascade mice models with *ex vivo* glycoreprogrammed T cells to evaluate the evolution of the disease and gain new insights on the potential of glycome switching in immunosuppression and cancer eradication. Furthermore, we could also explore the potential of kifunensine *in vivo*, by directly administering to these mice. In this case, we could evaluate a dual role of kifunensine, combining previous findings of our group (by increasing surface expression of high-mannose *N*-glycans on cancer cells) with this novel knowledge of T cell glycoreprogramming. To summarize, **Image 25** represents the graphical abstract regarding the work presented in this thesis.

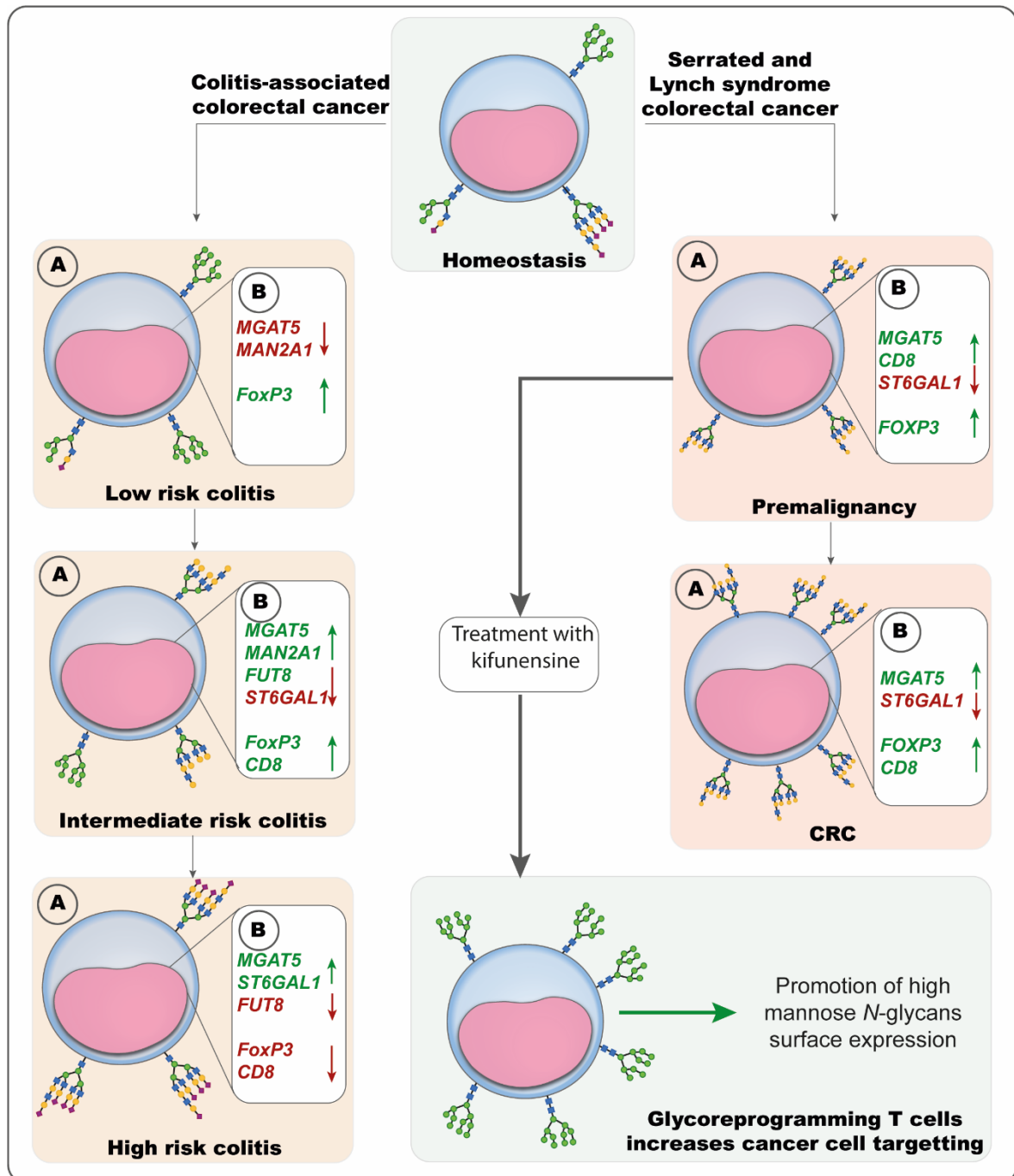


Figure 25 | T cell glycoproteins in CAC cascade (progression from inflammation to premalignancy and finally to carcinoma) and in serrated and Lynch syndrome cascades (progression from premalignancy to carcinoma). (A) T cell N-glycan surface expression along the cascades. (B) Glycogene and immune gene expression changes relatively to homeostatic T cell profile. The CAC cascade preliminary results suggest an increase in expression of β -1,6-linked branched N-glycans on T cell surface along the inflammatory pathway according to colitis risk. Moreover, this seems to be associated with an increase infiltration of immunosuppressive cells. However, we observed a trend decrease in infiltration of these Tregs in the stage which branched N-glycans seem to peak. Serrated and Lynch syndrome cascades are characterized by over expression of these branched structures, which are associated with decreased cancer cell targeting. These glycome changes occur early, as we demonstrated. Finally, we demonstrated that glycoreprogramming T cells has an immense potential as a directed therapy for CRC resolution.

7. References

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* 71 (3) (2021) 209–249.
2. Lortet-Tieulent, J., Georges, D., Bray, F. & Vaccarella, S. Profiling global cancer incidence and mortality by socioeconomic development. *Int. J. Cancer* 147 (11) (2020) 3029–3036.
3. Richards, M., Anderson, M., Carter, P., Ebert, B. L. & Mossialos, E. The impact of the COVID-19 pandemic on cancer care. *Nat. Cancer* 2020 16 1 (6) (2020) 565–567.
4. Balmain, A., Gray, J. & Ponder, B. The genetics and genomics of cancer. *Nat. Genet.* 33 (3S) (2003) 238–244.
5. Tomlinson, I. P. M., Novelli, M. R. & Bodmer, W. F. The mutation rate and cancer. *Proc. Natl. Acad. Sci. U. S. A.* 93 (25) (1996) 14800–14803.
6. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* 144 (5) (2011) 646–74.
7. Hanahan, D. & Coussens, L. M. Accessories to the Crime: Functions of Cells Recruited to the Tumor Microenvironment. *Cancer Cell* 21 (3) (2012) 309–322.
8. Fouad, Y. A. & Aanei, C. Revisiting the hallmarks of cancer. *Am. J. Cancer Res.* 7 (5) (2017) 1016–1036.
9. Arnold, M. *et al.* Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 66 (4) (2017) 683–691.
10. Favoriti, P. *et al.* Worldwide burden of colorectal cancer: a review. *Updates Surg.* 68 (1) (2016) 7–11.
11. Keum, N. N. & Giovannucci, E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat. Rev. Gastroenterol. Hepatol.* 16 (12) (2019) 713–732.
12. Nguyen, L. H., Goel, A. & Chung, D. C. Pathways of Colorectal Carcinogenesis. *Gastroenterology* 158 (2) (2020) 291–302.
13. Montminy, E. M., Jang, A., Conner, M. & Karlitz, J. J. Screening for Colorectal Cancer. *Med. Clin. North Am.* 104 (6) (2020) 1023–1036.

14. Xie, Y.-H., Chen, Y.-X. & Fang, J.-Y. Comprehensive review of targeted therapy for colorectal cancer. *Signal Transduct. Target. Ther.* 5 (1) (2020) 22.
15. Global Cancer Observatory - GLOBOCAN 2020. <https://gco.iarc.fr/today/>.
16. Pitchumoni, C. S. & Broder, A. Epidemiology of colorectal cancer. in *Colorectal Neoplasia and the Colorectal Microbiome* vol. 2 (Elsevier, 2020). 5–33.
17. Akimoto, N. *et al.* Rising incidence of early-onset colorectal cancer — a call to action. *Nat. Rev. Clin. Oncol.* 18 (4) (2021) 230–243.
18. Mundade, R., Imperiale, T. F., Prabhu, L., Loehrer, P. J. & Lu, T. Genetic pathways, prevention, and treatment of sporadic colorectal cancer. *Oncoscience* 1 (6) (2014) 400–406.
19. Ladabaum, U., Dominitz, J. A., Kahi, C. & Schoen, R. E. Strategies for Colorectal Cancer Screening. *Gastroenterology* 158 (2) (2020) 418–432.
20. Iwai, Y., Hamanishi, J., Chamoto, K. & Honjo, T. Cancer immunotherapies targeting the PD-1 signaling pathway. *J. Biomed. Sci.* 2017 241 24 (1) (2017) 1–11.
21. Lichtenstern, C. R., Ngu, R. K., Shalapour, S. & Karin, M. Immunotherapy, Inflammation and Colorectal Cancer. *Cells* 2020, Vol. 9, Page 618 9 (3) (2020) 618.
22. Conteduca, V., Sansonno, D., Russi, S. & Dammacco, F. Precancerous colorectal lesions (Review). *Int. J. Oncol.* 43 (4) (2013) 973–984.
23. Al-Sohaily, S., Biankin, A., Leong, R., Kohonen-Corish, M. & Warusavitarne, J. Molecular pathways in colorectal cancer. *J. Gastroenterol. Hepatol.* 27 (9) (2012) 1423–1431.
24. Nakanishi, Y., Diaz-Meco, M. T. & Moscat, J. Serrated Colorectal Cancer: The Road Less Travelled? *Trends in cancer* 5 (11) (2019) 742.
25. Liu, C. *et al.* Sessile serrated adenomas with dysplasia: morphological patterns and correlations with MLH1 immunohistochemistry. *Mod. Pathol.* 2017 3012 30 (12) (2017) 1728–1738.
26. Dias, A. M. *et al.* Glycans as critical regulators of gut immunity in homeostasis and disease. (2018) doi:10.1016/j.cellimm.2018.07.007.
27. Kaplan, G. G. & Ng, S. C. Understanding and Preventing the Global Increase of Inflammatory Bowel Disease. *Gastroenterology* 152 (2) (2017) 313-321.e2.

28. Guan, Q. A Comprehensive Review and Update on the Pathogenesis of Inflammatory Bowel Disease. *J. Immunol. Res.* 2019 **(2019)**.
29. Rogler, G. Chronic ulcerative colitis and colorectal cancer. *Cancer Lett.* 345 **(2) (2014)** 235–241.
30. Alatab, S. *et al.* The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol. Hepatol.* 5 **(1) (2020)** 17–30.
31. Thorsteinsdottir, S., Gudjonsson, T., Nielsen, O. H., Vainer, B. & Seidelin, J. B. Pathogenesis and biomarkers of carcinogenesis in ulcerative colitis. *Nat. Rev. Gastroenterol. Hepatol.* 8 **(7) (2011)** 395–404.
32. Scarpa, M. *et al.* Inflammatory colonic carcinogenesis: A review on pathogenesis and immunosurveillance mechanisms in ulcerative colitis. *World J. Gastroenterol.* 20 **(22) (2014)** 6774.
33. Al Bakir, I., Curtius, K. & Graham, T. A. From colitis to cancer: An evolutionary trajectory that merges maths and biology. *Frontiers in Immunology* vol. 9 **(OCT) (2018)** 2368.
34. Baker, A. M. *et al.* Evolutionary history of human colitis-associated colorectal cancer. *Gut* 68 **(6) (2019)** 985–995.
35. Kinugasa, T. & Akagi, Y. Status of colitis-associated cancer in ulcerative colitis. *World J. Gastrointest. Oncol.* 8 **(4) (2016)** 351.
36. Nicholson, L. B. The immune system. *Essays Biochem.* 60 **(3) (2016)** 275–301.
37. Hoebe, K., Janssen, E. & Beutler, B. The interface between innate and adaptive immunity. *Nat. Immunol.* 5 **(10) (2004)** 971–974.
38. Marshall, J. S., Warrington, R., Watson, W. & Kim, H. L. An introduction to immunology and immunopathology. *Allergy, Asthma Clin. Immunol.* 2018 142 14 **(2) (2018)** 1–10.
39. Theofilopoulos, A. N., Kono, D. H. & Baccala, R. The multiple pathways to autoimmunity. *Nat. Immunol.* 18 **(7) (2017)** 716–724.
40. Vinay, D. S. *et al.* Immune evasion in cancer: Mechanistic basis and therapeutic strategies. *Semin. Cancer Biol.* 35 **(2015)** S185–S198.

41. Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: From immunosurveillance to tumor escape. *Nat. Immunol.* 3 **(11)** **(2002)** 991–998.
42. Marshall, J. S., Warrington, R., Watson, W. & Kim, H. L. An introduction to immunology and immunopathology. *Allergy, Asthma Clin. Immunol.* 14 **(Suppl 2)** **(2018)**.
43. Choy, M. C., Visvanathan, K. & De Cruz, P. An Overview of the Innate and Adaptive Immune System in Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* 23 **(1)** **(2017)** 2–13.
44. Netea, M. G., Schlitzer, A., Placek, K., Joosten, L. A. B. & Schultze, J. L. Innate and Adaptive Immune Memory: an Evolutionary Continuum in the Host's Response to Pathogens. *Cell Host Microbe* 25 **(1)** **(2019)** 13–26.
45. Murphy, K. M. & Reiner, S. L. The lineage decisions of helper T cells. *Nat. Rev. Immunol.* 2 **(12)** **(2002)** 933–944.
46. Wan, Y. Y. Multi-tasking of helper T cells. *Immunology* 130 **(2)** **(2010)** 166–171.
47. Raphael, I., Nalawade, S., Eagar, T. N. & Forsthuber, T. G. T cell subsets and their signature cytokines in autoimmune and inflammatory diseases. *Cytokine* 74 **(1)** **(2015)** 5.
48. O'Connor, W., Zenewicz, L. A. & Flavell, R. A. The dual nature of T_H 17 cells: shifting the focus to function. *Nat. Immunol.* 2010 116 11 **(6)** **(2010)** 471–476.
49. Vignali, D. A. A., Collison, L. W. & Workman, C. J. How regulatory T cells work. *Nat. Rev. Immunol.* 8 **(7)** **(2008)** 523–532.
50. Farhood, B., Najafi, M. & Mortezaee, K. CD8⁺ cytotoxic T lymphocytes in cancer immunotherapy: A review. *J. Cell. Physiol.* 234 **(6)** **(2019)** 8509–8521.
51. Garber, K. $\gamma\delta$ T cells bring unconventional cancer-targeting to the clinic — again. *Nat. Biotechnol.* 38 **(4)** **(2020)** 389–391.
52. Krijgsman, D., Hokland, M. & Kuppen, P. J. K. The Role of Natural Killer T Cells in Cancer—A Phenotypical and Functional Approach. *Front. Immunol.* 0 **(FEB)** **(2018)** 367.
53. Dunn, G. P., Old, L. J. & Schreiber, R. D. The three Es of cancer immunoediting. *Annu. Rev. Immunol.* 22 **(4)** **(2004)** 329–360.

54. Bhatia, A. & Kumar, Y. Cancer-Immune Equilibrium: Questions Unanswered. *Cancer Microenviron.* 4 (2) (2011) 209.
55. Dong, P., Xiong, Y., Yue, J., Hanley, S. J. B. & Watari, H. Tumor-Intrinsic PD-L1 Signaling in Cancer Initiation, Development and Treatment: Beyond Immune Evasion. *Front. Oncol.* 0 (SEP) (2018) 386.
56. O'Donnell, J. S., Teng, M. W. L. & Smyth, M. J. Cancer immunoediting and resistance to T cell-based immunotherapy. *Nat. Rev. Clin. Oncol.* 16 (3) (2019) 151–167.
57. Kuol, N., Stojanovska, L., Nurgali, K. & Apostolopoulos, V. The mechanisms tumor cells utilize to evade the host's immune system. *Maturitas* 105 (2017) 8–15.
58. Muenst, S. *et al.* The immune system and cancer evasion strategies: Therapeutic concepts. *J. Intern. Med.* 279 (6) (2016) 541–562.
59. Hoos, A. Development of immuno-oncology drugs-from CTLA4 to PD1 to the next generations. *Nat. Rev. Drug Discov.* 15 (4) (2016) 235–247.
60. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* 15 (9) (2015) 540–555.
61. Nagae, M., Yamaguchi, Y., Taniguchi, N. & Kizuka, Y. 3D Structure and Function of Glycosyltransferases Involved in N-glycan Maturation. *Int. J. Mol. Sci.* 2020, Vol. 21, Page 437 21 (2) (2020) 437.
62. Reily, C., Stewart, T. J., Renfrow, M. B. & Novak, J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* 15 (6) (2019) 346–366.
63. Ohtsubo, K. & Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* 126 (5) (2006) 855–867.
64. Schjoldager, K. T., Narimatsu, Y., Joshi, H. J. & Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* 21 (12) (2020) 729–749.
65. Dias, A. M. *et al.* Dysregulation of T cell receptor N-glycosylation: A molecular mechanism involved in ulcerative colitis. *Hum. Mol. Genet.* 23 (9) (2014) 2416–2427.
66. Maverakis, E. *et al.* Glycans In The Immune system and The Altered Glycan Theory of Autoimmunity: A Critical Review. *J. Autoimmun.* 0 (2015) 1.
67. Lombard, J. The multiple evolutionary origins of the eukaryotic N- glycosylation pathway. *Biol. Direct* 2016 111 11 (1) (2016) 1–31.

68. Bieberich, E. Synthesis, Processing, and Function of N-glycans in N-glycoproteins. in *Advances in neurobiology* vol. 9 **(NIH Public Access, 2014)**. 47–70.
69. Stanley, P., Schachter, H. & Taniguchi, N. N-Glycans. *Cold Spring Harb. Lab. Press* **(2009)** 1–14.
70. Schachter, H. The joys of HexNAc. The synthesis and function of N- and O-glycan branches. *Glycoconj. J.* 17 **(7–9)** **(2000)** 465–483.
71. Pereira, M. S. *et al.* Glycans as key checkpoints of T cell activity and function. *Front. Immunol.* 9 **(NOV)** **(2018)** 1–13.
72. De Bousser, E., Meuris, L., Callewaert, N. & Festjens, N. Human T cell glycosylation and implications on immune therapy for cancer. *Hum. Vaccines Immunother.* 00 **(00)** **(2020)** 1–15.
73. Chien, M. W., Fu, S. H., Hsu, C. Y., Liu, Y. W. & Sytwu, H. K. The modulatory roles of N-glycans in t-cell-mediated autoimmune diseases. *Int. J. Mol. Sci.* 19 **(3)** **(2018)** 1–16.
74. Dias, A. M. *et al.* Metabolic control of T cell immune response through glycans in inflammatory bowel disease. *Proc. Natl. Acad. Sci. U. S. A.* 115 **(20)** **(2018)** E4651–E4660.
75. Carvalho, S. *et al.* Preventing E-cadherin aberrant N-glycosylation at Asn-554 improves its critical function in gastric cancer. *Oncogene* 35 **(13)** **(2016)** 1619–1631.
76. Pinho, S. S. *et al.* E-cadherin and adherens-junctions stability in gastric carcinoma: Functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and v. *Biochim. Biophys. Acta - Gen. Subj.* 1830 **(3)** **(2013)** 2690–2700.
77. Carvalho, S., Reis, C. A. & Pinho, S. S. Cadherins Glycans in Cancer: Sweet Players in a Bitter Process. *Trends in Cancer* 2 **(9)** **(2016)** 519–531.
78. Silva, M. C. *et al.* Glycans as immune checkpoints: removal of branched N-glycans enhances immune recognition preventing cancer progression. *Cancer Immunol. Res.* **(2020)** canimm.0264.2020 doi:10.1158/2326-6066.CIR-20-0264.
79. Magro, F. *et al.* Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders. *J. Crohn's Colitis* 11 **(6)** **(2017)** 649–670.

80. Cummings, R. D., Darvill, A. G., Etzler, M. E. & Hahn, M. G. Glycan-Recognizing Probes as Tools. *Essentials Glycobiol.* **(2017)** 611–625 doi:10.1101/GLYCOBIOLOGY.3E.048.
81. Chien, M. W., Fu, S. H., Hsu, C. Y., Liu, Y. W. & Sytwu, H. K. The modulatory roles of N-glycans in t-cell-mediated autoimmune diseases. *Int. J. Mol. Sci.* 19 **(3)** **(2018)** 1–11.
82. Rabinovich, G. A., Kooyk, Y. van & Cobb, B. A. Glycobiology of immune responses. *Ann. N. Y. Acad. Sci.* 1253 **(1)** **(2012)** 1–15.
83. Grigorian, A. & Demetriou, M. Mgat5 Deficiency in T Cells and Experimental Autoimmune Encephalomyelitis . *ISRN Neurol.* 2011 **(2011)** 1–6.
84. Demetriou, M., Granovsky, M., Quaggin, S. & Dennis, J. W. Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation. *Nature* 409 **(6821)** **(2001)** 733–739.
85. Idos, G. E. *et al.* The Prognostic Implications of Tumor Infiltrating Lymphocytes in Colorectal Cancer: A Systematic Review and Meta-Analysis. *Sci. Reports* 2020 101 10 **(1)** **(2020)** 1–14.
86. Buccitelli, C. & Selbach, M. mRNAs, proteins and the emerging principles of gene expression control. *Nat. Rev. Genet.* 21 **(10)** **(2020)** 630–644.
87. Cedeno-Laurent, F. & Dimitroff, C. J. Galectin-1 research in T cell immunity: Past, present and future. *Clin. Immunol.* 142 **(2)** **(2012)** 107.
88. Bissahoyo, A. *et al.* Azoxymethane is a genetic background-dependent colorectal tumor initiator and promoter in mice: Effects of dose, route, and diet. *Toxicol. Sci.* 88 **(2)** **(2005)** 340–345.
89. Méndez-Huergo, S. P., Blidner, A. G. & Rabinovich, G. A. Galectins: emerging regulatory checkpoints linking tumor immunity and angiogenesis. *Curr. Opin. Immunol.* 45 **(2017)** 8–15.
90. Elbein, A. D., Tropea, J. E., Mitchell, M. & Kaushal, G. P. Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I. *J. Biol. Chem.* 265 **(26)** **(1990)** 15599–15605.
91. Grigorian, A. *et al.* Control of T cell-mediated autoimmunity by metabolite flux to N-glycan biosynthesis. *J. Biol. Chem.* 282 **(27)** **(2007)** 20027–20035.
92. Lalos, A. *et al.* Prognostic significance of CD8+ T-cells density in stage III colorectal

- cancer depends on SDF-1 expression. *Sci. Reports* 2021 111 11 (1) (2021) 1–11.
93. Smith, L. K. *et al.* Interleukin-10 Directly Inhibits CD8+ T Cell Function by Enhancing N-Glycan Branching to Decrease Antigen Sensitivity. *Immunity* 48 (2) (2018) 299–312.e5.
 94. Cagnoni, A. J. *et al.* Galectin-1 fosters an immunosuppressive microenvironment in colorectal cancer by reprogramming CD8 + regulatory T cells. doi:10.1073/pnas.2102950118/-/DCSupplemental.
 95. Nielsen, M. I. *et al.* Galectin binding to cells and glycoproteins with genetically modified glycosylation reveals galectin–glycan specificities in a natural context. *J. Biol. Chem.* 293 (52) (2018) 20249–20262.
 96. Sundblad, V., Morosi, L. G., Geffner, J. R. & Rabinovich, G. A. Galectin-1: A Jack-of-All-Trades in the Resolution of Acute and Chronic Inflammation. *J. Immunol.* 199 (11) (2017) 3721–3730.
 97. Toscano, M. A. *et al.* Differential glycosylation of TH1, TH2 and TH-17 effector cells selectively regulates susceptibility to cell death. *Nat. Immunol.* 8 (8) (2007) 825–834.
 98. Shi, X. *et al.* Targeting glycosylation of PD-1 to enhance CAR-T cell cytotoxicity. *J. Hematol. Oncol.* 2019 121 12 (1) (2019) 1–4.
 99. Li, C.-W. *et al.* Glycosylation and stabilization of programmed death ligand-1 suppresses T-cell activity. *Nat. Commun.* 2016 71 7 (1) (2016) 1–11.
 100. Sun, R., Kim, A. M. J. & Lim, S. O. Glycosylation of immune receptors in cancer. *Cells* 10 (5) (2021).
 101. Cabral, J. *et al.* Distinctive Surface Glycosylation Patterns Associated With Mouse and Human CD4+ Regulatory T Cells and Their Suppressive Function. *Front. Immunol.* 8 (AUG) (2017).
 102. Verhelst, X. *et al.* Protein Glycosylation as a Diagnostic and Prognostic Marker of Chronic Inflammatory Gastrointestinal and Liver Diseases. *Gastroenterology* 158 (1) (2020) 95–110.
 103. Goss, P. E., Baptiste, J., Fernandes, B., Baker, M. & Dennis, J. W. A phase I study of swainsonine in patients with advanced malignancies. *Cancer Res.* 54 (6) (1994) 1450–7.
 104. Baptista, J. A. *et al.* Measuring swainsonine in serum of cancer patients: phase I clinical trial. *Clin. Chem.* 40 (3) (1994) 426–430.

105. Goss, P. E., Reid, C. L., Bailey, D. & Dennis, J. W. Phase IB clinical trial of the oligosaccharide processing inhibitor swainsonine in patients with advanced malignancies. *Clin. Cancer Res.* 3 (7) **(1997)** 1077–86.
106. Mohla, S. *et al.* Swainsonine: a new antineoplastic immunomodulator. *J. Natl. Med. Assoc.* 81 (10) **(1989)** 1049.