



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia**

JÉSSICA DE ANDRADE DA COSTA

**Avaliação de glicobiomarcadores em células tronco tumorais de câncer
colorretal**

Orientador: Prof. Dr. Julio Cesar Madureira de Freitas Junior

**RIO DE JANEIRO
2020**



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia**

JÉSSICA DE ANDRADE DA COSTA

Avaliação de glicobiomarcadores em células tronco tumorais de câncer colorretal

Dissertação apresentada ao Instituto Nacional de Câncer como parte dos requisitos para obtenção do título de Mestre em Oncologia.

Orientador: Prof. Dr. Julio Cesar Madureira de Freitas Junior

**RIO DE JANEIRO
2020**

C837a Costa, Jéssica de Andrade
Avaliação de glicobiomarcadores em células tronco tumorais de câncer colorretal. / Jéssica de Andrade Costa – Rio de Janeiro, 2020.
171 : il. Color.

Dissertação (Mestrado em Oncologia) – Programa de Pós-Graduação em Oncologia, Instituto Nacional de Câncer José Alencar Gomes da Silva, 2020.

Orientador: Prof. Dr. Júlio Cesar Madureira de Freitas Junior.

1. Neoplasias Colorretais. 2. Células-Tronco. 3. Polissacarídeos. I. Freitas Junior, Júlio Cesar Madureira. (Orient.). II. Instituto Nacional de Câncer José Alencar Gomes da Silva.

III. Título.

CDD 616.994347042

Catálogo na fonte
Kátia Simões CRB7/5952



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia**

JÉSSICA DE ANDRADE DA COSTA

**AVALIAÇÃO DE GLICOBIMARCADORES EM CÉLULAS TRONCO TUMORAIS
DE CÂNCER COLORRETAL**

ORIENTADOR: Prof. Dr. Julio Cesar Madureira de Freitas Junior

Dissertação aprovada em: 24 / 03 / 2020

EXAMINADORES:

**Prof. Dra. Nathalia de Oliveira Meireles da Costa – INCA
Prof. Dra. Stephany Cristiane Corrêa – INCA
Prof. Dra. Adriane Regina Todeschini – UFRJ
Prof. Dra. Patrícia Abrão Possik - INCA – Suplente I
Prof. Dr. Frederico Alisson da Silva – UFRJ – Suplente II**

**RIO DE JANEIRO
2020**

DEDICATÓRIA

À minha querida tia Áurea que, embora já tenha partido, permanece sempre viva em meu coração e em minhas memórias da infância.

AGRADECIMENTOS

À Deus por ser minha fortaleza durante os momentos mais difíceis do projeto e por sempre operar pequenos milagres diários na minha vida.

À minha mãe, Maria do Socorro, e meu pai, Antonio Cesar, pelo amor, apoio incondicional mesmo durante meus piores dias e por me ensinarem os valores mais importantes da vida.

À minha irmã Samara por ser a minha psicóloga, conselheira e cúmplice.

À minha madrinha Helena e meu padrinho Donald pelo constante apoio, carinho e demonstração de orgulho pelas minhas conquistas.

Ao Dr. José Morgado por ter me acolhido em seu laboratório quando ainda dava meus primeiros passos na carreira científica, isso foi crucial para meu crescimento e amadurecimento profissional.

Ao meu orientador e amigo Julio pela paciência, conselhos, discussões construtivas, ensinamentos, apoio nos momentos de incerteza e por ter me apresentado esse universo fascinante da Glicobiologia.

À Sheila por ter embarcado na ideia do projeto, ter dispensado seu tempo ajudando ao longo de toda a construção do modelo e por ter me contagiado a cada reunião de dados com seu entusiasmo de ver resultados que eu considerava pequenos como grandes passos dentro do projeto.

À Pat pelas longas discussões a respeito das células tronco e por estar sempre disposta a ajudar.

Ao Léo Karam por todos os ensinamentos a respeito de RT-qPCR e por se mostrar sempre disposto a contribuir seja com protocolos ou reagentes.

À Marina e à Paula por toda a paciência e assessoria na busca por reagentes.

Às *glico girls* Érika, Michelle e Nathália por todas as risadas, momentos de descontração e por serem mais do que apenas colegas de trabalho.

Aos integrantes do Programa de Oncobiologia Celular e Molecular, Cássio, Renata, Adriana, Murilo, Clara, Bruna, Dani, Gabi, Mari, Luciana, Ana Emília que dividiram bancadas e momentos ao meu lado. Em especial, à Annie por me salvar sempre quando o assunto era RT-qPCR, por ouvir minhas loucuras seja com relação à carreira ou à vida e por me aconselhar nos momentos de dificuldade. À Josi por tornar os dias no lab sempre mais leves com sua alegria contagiante e seu enorme coração. À Deiró por todo o carinho durante meu processo seletivo do mestrado. Nunca esquecerei aquele abraço de mãe. Obrigada!

Aos membros antigos do lab que passaram pela minha vida deixando suas marcas. Ao Wal por nunca ter me enganado sobre o que era a pesquisa com sua frase de efeito “a vida é dura”, ao Wallace por me ensinar a pensar criticamente e por ser um exemplo de que ninguém é capaz de parar um talento, ao Perony por todas as conversas malucas e engraçadas, à Bruna por todos os conselhos, à Taíssa pelas palavras certas nos momentos de aflição e por sempre acreditar mais em mim do que eu mesma, à Jenny e Andrea pelo companheirismo nos meus primeiros anos de laboratório, à Pri, Natália, Rubem e Obina que me ensinaram muito do que sei hoje, à Lili pelo exemplo de determinação e, em especial, ao Pedro Barcellos por ter sido peça fundamental para a continuidade desse projeto no momento em que os recursos foram mais escassos.

Aos amigos que conheci no INCA. Ao J. Show por todas as piadas na copa que sempre aliviaram as pressões do dia a dia, à Isa pelas discussões científicas e em especial ao Renan pela forma atenta de reparar em cada detalhe das pessoas e de se doar pelos outros.



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

AVALIAÇÃO DE GLICOBIMARCADORES EM CÉLULAS TRONCO TUMORAIS DE CÂNCER COLORRETAL

RESUMO

DISSERTAÇÃO DE MESTRADO

Jéssica de Andrade da Costa

O câncer colorretal (CCR) apresenta alta taxa de mortalidade, resultado dos processos de metástase e recorrência da doença. Acredita-se que as células tronco tumorais (CTTs) são cruciais para ambos os processos, pois asseguram a manutenção da massa tumoral, além de serem intrinsecamente resistentes às terapias convencionais. Embora diversos trabalhos venham tentando caracterizar as CTTs, aqueles que se dedicam a avaliar o papel de glicanos (carboidratos covalentemente ligados a proteínas ou lipídios) na identificação dessa população e na manutenção do fenótipo tronco ainda são relativamente poucos. Assim, o objetivo do presente estudo foi investigar e caracterizar glicobimarcadores em subpopulações de células tronco tumorais de câncer colorretal. Para isso, inicialmente foi padronizado um protocolo de formação de esferas para as linhagens Caco-2 e HT-29. Após o crescimento das esferas, foram monitoradas variações na expressão de cinco conhecidos marcadores de CTTs (*CD24*, *CD44*, *ALDH1*, *LGR5* e *PROM1*). Em comparação ao que se observa na condição parental (2D), em Caco-2 foi vista uma redução da expressão de *CD24*, enquanto em HT-29 foi observado um aumento dos níveis de expressão de *ALDH1*, *LGR5* e *PROM1*. Com relação aos glicogenes (genes relevantes para a glicosilação), oito destes tiveram seus níveis de expressão monitorados (*ST3GAL1*, *OGT*, *OGA*, *MGAT5*, *GFAT1*, *GFAT2*, *B4GALT1* e *B3GNT2*), tendo sido observado aumento de *B3GNT2*, *OGT* e *OGA* na condição de esferas em HT-29. Já em Caco-2, amplas variações experimentais impossibilitaram qualquer conclusão no que tange a mudanças na expressão de glicogenes. Análises de correlação *in silico* (marcadores de CTTs versus glicogenes), utilizando amostras de carcinoma de cólon e reto provenientes do *The Cancer Genome Atlas Program*, mostraram uma correlação positiva entre a expressão de *CD44xOGA*, *CD44xB3GNT2* e *LGR5xOGA* para carcinomas de cólon e *CD44xOGA*, *LGR5xOGA*, *CD44xOGT* e *CD24xB3GNT2* para carcinomas de reto. Também foi observado aumento na expressão de *LGR5*, *CD44*, *MGAT5*, *OGA*, *OGT* e *ST3GAL1* e diminuição de *CD24*, *ALDH1* e *B4GALT1* em amostras de carcinoma de cólon quando comparadas ao tecido normal adjacente, bem como aumento na expressão de *LGR5*, *CD44*, *PROM1* e *OGT* e redução de *ALDH1*, *B3GNT2* e *ST3GAL1* em amostras de carcinoma de reto quando comparadas com amostras de tecido normal adjacente. Surpreendentemente, nenhum dos marcadores de CTTs foi capaz de prever a sobrevida livre de doença (SLD) ou a sobrevida global (SG) de pacientes com CCR, no entanto, pacientes com carcinoma de cólon que apresentaram maior expressão de *MGAT5* tiveram menor SLD. Além disso, a maior expressão de *OGA* também resultou em menor SG de pacientes com câncer de cólon. Portanto, nossos dados sugerem que os níveis de expressão *OGA*, além de correlacionarem-se com os de marcadores de CTTs, impactam diretamente na sobrevida de pacientes com carcinoma de cólon, fornecendo indícios de uma possível função dessa enzima em mecanismos de manutenção do fenótipo tronco.



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

ASSESSMENT OF GLYCOBIOMARKERS IN COLORECTAL CANCER STEM CELLS

ABSTRACT

MASTER THESIS

Jéssica de Andrade da Costa

Colorectal cancer (CRC) has a high mortality rate, resulting from the processes of metastasis and disease recurrence. Cancer stem cells (CSCs) are believed to be crucial for both processes, as they ensure the maintenance of tumor bulk, in addition to being intrinsically resistant to conventional therapies. Although several studies have been trying to characterize CSCs, those dedicated to assessing the role of glycans (carbohydrates covalently linked to proteins or lipids) in the identification of this population and in the maintenance of stemness are yet relatively scarce. Thus, the aim of the present study was to investigate and characterize glyco-biomarkers in colorectal cancer stem cell subpopulations. For this purpose, a sphere formation assay was initially standardized for Caco-2 and HT-29 cell lines. After the spheres growth, variations in the expression of five known CSC markers (*CD24*, *CD44*, *ALDH1*, *LGR5* e *PROM1*) were monitored. In comparison to the parental condition (2D), a reduction of *CD24* expression was seen in Caco-2, while in HT-29 an increase in the expression levels of *ALDH1*, *LGR5* and *PROM1* was observed. Regarding glycosylated genes (glycosylation relevant genes), eight of them (*ST3GAL1*, *OGT*, *OGA*, *MGAT5*, *GFAT1*, *GFAT2*, *B4GALT1* e *B3GNT2*) have had their expression monitored. An increase in *B3GNT2*, *OGT* and *OGA* was observed in HT-29 sphere condition. In Caco-2, on the other hand, wide experimental variations precluded any conclusion regarding changes in glycosylated genes expression. *In silico* correlation analyzes (CSCs markers versus glycosylated genes) using colon and rectum carcinoma samples from The Cancer Genome Atlas Program showed a positive correlation between *CD44xOGA*, *CD44xB3GNT2* and *LGR5xOGA* expressions for colon carcinoma and *CD44xOGA*, *LGR5xOGA*, *CD44xOGT* and *CD24xB3GNT2* for rectum carcinoma. In addition, increased *LGR5*, *CD44*, *MGAT5*, *OGA*, *OGT* and *ST3GAL1* expressions and decreased *CD24*, *ALDH1* and *B4GALT1* expressions were also observed in colon carcinoma samples when compared to adjacent normal tissue. Moreover, increased *LGR5*, *CD44*, *PROM1* and *OGT* expressions and reduced *ALDH1*, *B3GNT2* and *ST3GAL1* expressions were noted in rectal carcinoma samples when compared to adjacent normal tissue samples. Surprisingly, none of the CSC markers was able to predict disease-free survival (DFS) or the overall survival (OS) of CRC patients, however, colon carcinoma patients who have higher *MGAT5* expression had lower DFS. In addition, greater *OGA* expression resulted in lower OS of colon carcinoma patients. Therefore, our data suggest that *OGA* expression correlates with CSC markers and directly impacts the survival of colon carcinoma patients suggesting a possible function for this enzyme in the regulation of stemness.

LISTA DE ILUSTRAÇÕES

Figura 1.1: Distribuição proporcional por sexo dos dez tipos de câncer mais incidentes no Brasil, segundo estimativas para 2020, com exceção de pele não melanoma.....	1
Figura 1.2: Distribuição do CCR de acordo com as manifestações esporádica, familiar e hereditária.....	4
Figura 1.3: Subtipos moleculares do CCR segundo o classificador CMS.....	7
Figura 1.4: Tipos de heterogeneidade intratumoral.....	9
Figura 1.5: Modelo de funcionamento das células tronco em condições normais e tumorais.....	13
Figura 1.6: Representação das diferentes classes de glicoconjugados presentes nas células.....	17
Figura 1.7: Principais tipos de <i>N</i> -glicanos.....	18
Figura 1.8: Síntese de UDP-GlcNAc através da via das hexosaminas.....	19
Figura 1.9: Principais estruturas de glicanos relevantes para o câncer.....	22
Figura 5.1: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (3×10^3 células/poço) em meio M1.....	37
Figura 5.2: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (3×10^3 células/poço) em meio M2.....	38
Figura 5.3: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (3×10^3 células/poço) em meio M3.....	39
Figura 5.4: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (4×10^3 células/poço) em meio M1.....	40
Figura 5.5: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (4×10^3 células/poço) em meio M2.....	41
Figura 5.6: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (4×10^3 células/poço) em meio M3.....	42
Figura 5.7: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (4×10^3 células/poço) em meio M1 seguindo o protocolo padronizado.....	42
Figura 5.8: Acompanhamento da formação de esferas a partir do plaqueamento das células Caco-2 (4×10^3 células/poço) em meio M1 seguindo o protocolo padronizado.....	43

Figura 5.9: Variações na expressão dos marcadores de CTTs em Caco-2 e HT-29 após o ensaio de formação de esferas	45
Figura 5.10: Variações na expressão de glicogenes em Caco-2 e HT-29 submetidas ao ensaio de formação de esferas	46
Figura 5.11: Correlação linear entre a expressão de marcadores de CTTs e glicogenes em amostras de carcinoma de cólon.....	48
Figura 5.12: Correlação linear entre a expressão de marcadores de CTTs e glicogenes em amostras de carcinoma de reto	49
Figura 5.13: Diferenças na expressão dos marcadores de CTTs entre amostras normais adjacentes e de tumor primário de cólon.....	50
Figura 5.14: Diferenças na expressão dos marcadores de CTTs entre amostras normais adjacentes e de tumor primário de reto	51
Figura 5.15: Diferenças na expressão de marcadores de CTTs entre os estadios do CCR	52
Figura 5.16: Diferenças na expressão de marcadores de CTTs entre os estadios do CCR	53
Figura 5.17: Diferenças na expressão de glicogenes entre amostras normais adjacentes e de tumor primário de cólon	55
Figura 5.18: Diferenças na expressão de glicogenes entre amostras normais adjacentes e de tumor primário de reto.....	56
Figura 5.19: Diferenças na expressão de glicogenes entre os estadios do CCR.....	58
Figura 5.20: Diferenças na expressão de glicogenes entre os estadios do CCR.....	59
Figura 5.21: Expressão de <i>B4GALT1</i> e <i>OGA</i> de acordo com o acometimento linfonodal.....	60
Figura 5.22: Sobrevida global de acordo com os níveis de expressão de diferentes marcadores de CTTs em amostras de câncer de cólon.....	62
Figura 5.23: Sobrevida livre de doença de acordo com os níveis de expressão de diferentes marcadores de CTTs em amostras de câncer de cólon	63
Figura 5.24: Sobrevida global de acordo com os níveis de expressão de diferentes marcadores de CTTs em amostras de câncer de reto	64
Figura 5.25: Sobrevida livre de doença de acordo com os níveis de expressão de diferentes marcadores de CTTs em amostras de câncer de reto.....	65
Figura 5.26: Sobrevida global de acordo com os níveis de expressão de diferentes glicogenes em amostras de câncer de cólon	68

Figura 5.27: Sobrevida livre de doença de acordo com os níveis de expressão de diferentes glicogenes em amostras de câncer de cólon.....	70
Figura 5.28: Sobrevida livre de doença e sobrevida global de acordo com os níveis de expressão de <i>MGAT5</i> e <i>OGA</i> apenas de pacientes que não apresentam instabilidade de microssatélites.....	70

LISTA DE TABELAS

Tabela 4.1: Número de repetições em cada alelo da amostra de Caco-2 analisada e da referência (ATCC) para cada um dos sítios genéticos avaliados.	28
Tabela 4.2: Número de repetições em cada alelo da amostra de HT-29 analisada e da referência (ATCC) para cada um dos sítios genéticos avaliados.	29
Tabela 4.3: Lista contendo as sequências iniciadoras utilizadas nas análises de qPCR.	32
Tabela 5.1: Diferença entre as médias aritméticas dos Cqs de genes-referência entre as condições parental e esfera.	44
Tabela 5.2: Correlações lineares analisadas entre a expressão de marcadores de CTTs e glicogenes em amostras de câncer de cólon indicando o coeficiente de correlação e o valor de P	47
Tabela 5.3: Correlações lineares analisadas entre a expressão de marcadores de CTTs e glicogenes em amostras de câncer de reto indicando o coeficiente de correlação e o valor de P	47

LISTA DE ABREVIATURAS E SIGLAS

- 5-FU – 5-Fluorouracil
- ABC – do inglês *ATP-Binding Cassette*
- ALDH1 – do inglês *Aldehyde dehydrogenase 1*
- APC – do inglês *Adenomatous Polyposis Coli*
- BAG – benzil-2-acetamido-2-deoxi- α -D-galactopiranosídeo
- BLAST – do inglês *Basic Local Alignment Search Tool*
- BRAF – do inglês *v-raf murine sarcoma viral oncogene homolog B*
- BSA – do inglês *Bovine Serum Albumin*
- CAFs – do inglês *Cancer Associated Fibroblasts*
- CCR – Câncer colorretal
- CIMP – do inglês *CpG Island Methylator Phenotype*
- CIN – do inglês *Chromosomal Instability*
- CK20 – do inglês *Cytokeratin 20*
- CMP – citidina monofosfato
- Cq – do inglês *Quantification Cycle*
- CRIS – do inglês *Colorectal Cancer Intrinsic Subtypes*
- CTC – Células Tumorais Circulantes
- CTTs – Células Tronco Tumorais
- DEPC – dietilpirocarbonato
- DMEM – do inglês *Dulbecco's Modified Eagle Medium*
- EMT – do inglês *Epithelial Mesenchymal Transition*
- ESCs – do inglês *Embryonic Stem Cells*
- FABP2 – do inglês *Fatty Acid Binding Protein 2*
- FAP – do inglês *Familial Adenomatous Polyposis*
- Fuc – fucose
- GAG – Glicosaminoglicanos
- Gal – galactose
- GalNAc – N-acetilgalactosamina
- GDP – guanosina difosfato
- GFAT – glutamina-frutose-6-fosfato amidotransferase
- Glc – glicose
- GlcA – ácido glicurônico
- GlcNAc – N-acetilglicosamina

GPI – do inglês *Glycosylphosphatidylinositol*
HA – do inglês *Hyaluronic Acid*
HDAC6 – do inglês *Histone deacetylase 6*
HNPCC – do inglês *Hereditary Non-Polyposis Colorectal Cancer*
HR – do inglês *Hazard Ratio*
IDH – Índice de Desenvolvimento Humano
IGF2 – do inglês *Insulin-like Growth Factor 2*
iPS – do inglês *Induced Pluripotent Stem Cell*
JAK/STAT – do inglês *Janus Kinase/Signal Transducers and Activators of Transcription*
KRAS – do inglês *Kirsten rat sarcoma viral oncogene homolog*
LGR5 – do inglês *Leucine Rich Repeat Containing G Protein-Coupled Receptor 5*
Man – manose
MLH1 – do inglês *MutL Homolog 1*
MMR – do inglês *Mismatch Repair*
MSH2 – do inglês *MutS Homolog 2*
MSH6 – do inglês *MutS Homolog 6*
MSI – do inglês *Microsatellite Instability*
MUC2 – do inglês *Mucin 2*
Neu5Ac – ácido siálico
NF-κB – do inglês *Nuclear Factor-Kappa B*
OCT4 – do inglês *Octamer-binding transcription factor 4*
OGA – O-GlcNAcase
OGT – O-GlcNAc transferase
PDX – do inglês *Patient-Derived Xenografts*
PMS2 – do inglês *Postmeiotic Segregation Increased 2*
PROM1 – do inglês *Prominin 1*
PIK3CA – do inglês *Phosphatidylinositol 4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha*
SFB – Soro Fetal Bovino
SG – Sobrevida Global
SLD – Sobrevida Livre de Doença
SMAD4 – do inglês *SMAD family member 4*
SOX2 – do inglês *Sex-determining region Y-box 2*
SRC – do inglês *v-src avian sarcoma viral oncogene*

sTn – antígeno sialil Tn

STRs – do inglês *Short Tandem Repeats*

TCF/LEF – do inglês *T-cell factor/lymphoid enhancer binding factor*

TCGA – do inglês *The Cancer Genome Atlas*

TGF- β – do inglês *Transforming Growth Factor- β*

TNF- α – do inglês *Tumor Necrosis Factor α*

TP53 – do inglês *Tumor Protein p53*

UDP – uridina difosfato

VEGF – do inglês *Vascular Endothelial Growth Factor*

WGS – do inglês *Whole Genome Sequencing*

WNT – do inglês *wingless-type MMTV integration site family*

Xyl – xilose

SUMÁRIO

1. Introdução	1
1.1. Câncer colorretal	1
1.1.1. Epidemiologia e fatores de risco	1
1.1.2. Bases moleculares e heterogeneidade tumoral	4
1.2. Células tronco tumorais	11
1.2.1. Mecanismos de detecção e isolamento de células tronco	14
1.3. Glicosilação de proteínas	16
1.4. Papel dos glicanos em câncer	19
1.5. Papel dos glicanos no estado tronco tumoral	23
2. Justificativa do estudo	26
3. Objetivo	27
3.1. Objetivo Geral	27
3.2. Objetivos Específicos	27
4. Materiais e métodos	28
4.1. Cultura de células	28
4.2. Ensaio de formação de esferas	29
4.2.1. Composição dos diferentes meios testados	29
4.3. Análise da expressão gênica por RT-qPCR	30
4.3.1. Extração e quantificação de RNA	30
4.3.2. Síntese de DNA complementar (c-DNA)	31
4.3.3. Obtenção das sequências iniciadoras	31
4.3.4. Reação de PCR quantitativo em tempo real (qPCR)	33
4.4. Análises in silico	34
4.4.1. Análise comparativa da expressão gênica entre amostras normais e tumorais	34
4.4.2. Análises de correlação	34
4.4.3. Análises do desfecho clínico	35
4.5. Análise estatística	35
5. Resultados	36
5.1. Estabelecimento do ensaio de formação de esferas para enriquecimento de CTTs.	36
5.2. Identificação dos melhores genes-referência para as análises por RT-qPCR.	43

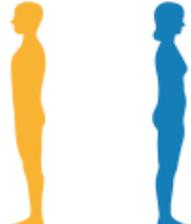
5.3. O ensaio de formação de esferas é capaz de induzir mudanças na expressão gênica de marcadores de CTTs em linhagens celulares de câncer colorretal.	44
5.4. Efeito do ensaio de formação de esferas na expressão de glicogenes em linhagens celulares de câncer colorretal.	45
5.5. Investigação <i>in silico</i> da correlação linear entre a expressão gênica dos marcadores de CTTs e a expressão de glicogenes.	46
5.6. Análises <i>in silico</i> do perfil de expressão gênica de marcadores de CTTs.	49
5.7. Análises <i>in silico</i> do perfil de expressão gênica de glicogenes.....	54
5.8. Avaliação do impacto da expressão de marcadores de CTTs sobre a sobrevida de pacientes com câncer colorretal	61
5.9. Avaliação do impacto da expressão de glicogenes sobre a sobrevida de pacientes com câncer colorretal.....	66
6. Discussão.....	72
7. Conclusão.....	85
8. Referências bibliográficas	86
9. Anexos.....	100

1. Introdução

1.1. Câncer colorretal

1.1.1. Epidemiologia e fatores de risco

O câncer figura entre os mais relevantes problemas de saúde pública, seja no Brasil ou no exterior. De acordo com dados da Organização Mundial da Saúde, uma em cada seis mortes ocorre em função do câncer, conferindo a ele o posto de segunda maior causa de morte no mundo (*World Health Organization, 2018*). Estimativas da GLOBOCAN (base de dados da Agência Internacional de Pesquisa em Câncer que reúne informações de 185 países) apontaram a ocorrência de 18,1 milhões de novos casos de câncer e 9,6 milhões de mortes decorrentes da doença para o ano de 2018 (BRAY *et al.*, 2018). No Brasil, a estimativa foi de aproximadamente 625 mil novos casos a cada ano do triênio 2020-2022 (MINISTÉRIO DA SAÚDE, 2019). O câncer colorretal (CCR) é um dos principais tipos de câncer que acometem a população mundial, ocupando a terceira posição em incidência e a segunda em mortalidade (BRAY *et al.*, 2018). No que diz respeito à população brasileira, quando se exclui o câncer de pele não melanoma, ele corresponde ao segundo tipo mais frequente em homens e mulheres, conforme ilustrado na figura 1.1 (MINISTÉRIO DA SAÚDE, 2019).

Localização Primária	Casos	%			Localização Primária	Casos	%	
Próstata	65.840	29,2%		Homens	Mama feminina	66.280	29,7%	
Cólon e reto	20.520	9,1%			Mulheres	Cólon e reto	20.470	9,2%
Traqueia, brônquio e pulmão	17.760	7,9%			Colo do útero	16.590	7,4%	
Estômago	13.360	5,9%			Traqueia, brônquio e pulmão	12.440	5,6%	
Cavidade oral	11.180	5,0%			Glândula tireoide	11.950	5,4%	
Esôfago	8.690	3,9%			Estômago	7.870	3,5%	
Bexiga	7.590	3,4%			Ovário	6.650	3,0%	
Linfoma não Hodgkin	6.580	2,9%			Corpo do útero	6.540	2,9%	
Laringe	6.470	2,9%			Linfoma não Hodgkin	5.450	2,4%	
Leucemias	5.920	2,6%			Sistema nervoso central	5.220	2,3%	

*Números arredondados para múltiplos de 10.

Figura 1.1: Distribuição proporcional por sexo dos dez tipos de câncer mais incidentes no Brasil, segundo estimativas para 2020, com exceção de pele não melanoma. Retirado de Estimativa 2020: incidência de câncer no Brasil. INCA-RJ.

O recente desenvolvimento econômico do Brasil repercute no aumento da incidência do CCR, uma vez que, para esse tipo de câncer, as taxas de incidência

são influenciadas pelo Índice de Desenvolvimento Humano (IDH) do país (FIDLER; SOERJOMATARAM; BRAY, 2016). Muito desse aumento é explicado pela incorporação de hábitos ocidentais que afetam tanto a dieta quanto o estilo de vida. No entanto, o que se observa em países altamente desenvolvidos é a redução da incidência e mortalidade devido à implementação de estratégias que visam a detecção precoce da doença, visto que a taxa de sobrevivência em cinco anos para pacientes diagnosticados nos estádios I e II é, respectivamente, de 91% e 82%, valores que podem chegar a apenas 12% para pacientes no estágio IV (ARNOLD *et al.*, 2017; MILLER *et al.*, 2019).

O desenvolvimento do CCR está atrelado não apenas a fatores genéticos, mas também a fatores ambientais. Um dos principais fatores de risco associados ao seu desenvolvimento é o envelhecimento. Além dele, também figuram nessa lista o tipo de dieta (rica em carne vermelha e alimentos embutidos, com alto consumo de gordura e baixo consumo de fibras), o sedentarismo, a obesidade, o tabagismo e o consumo de álcool (ARAN *et al.*, 2016; THANIKACHALAM; KHAN, 2019). Também são incluídos como fatores de risco o histórico familiar e as doenças inflamatórias que acometem o intestino, como a colite ulcerativa e a doença de Crohn (STIDHAM; HIGGINS, 2018; SCHUBERT *et al.*, 2019).

No que diz respeito às disfunções genéticas, a Síndrome de Lynch (ou câncer colorretal não poliposo hereditário; HNPCC, do inglês *Hereditary Non-Polyposis Colorectal Cancer*) e a Polipose Adenomatosa Familiar (FAP, do inglês *Familial Adenomatous Polyposis*) representam os fatores de risco mais comuns e mais bem relatados. A Síndrome de Lynch se caracteriza por se originar a partir de uma mutação de caráter autossômico dominante na linhagem germinativa, que afeta os genes *MLH1* (do inglês *MutL Homolog 1*), *MSH2* (do inglês *MutS Homolog 2*), *MSH6* (do inglês *MutS Homolog 6*) ou *PMS2* (do inglês *Postmeiotic Segregation Increased 2*), responsáveis pelo reparo do DNA durante o pareamento incorreto (MMR, do inglês *Mismatch Repair*). Como consequência, os pacientes acometidos por essa síndrome apresentam alta instabilidade genômica em pequenas regiões do DNA que apresentam de um a quatro pares de base repetidos em *tandem*, denominadas microssatélites (MA *et al.*, 2018). Já a FAP ocorre mediante uma mutação germinativa no gene supressor tumoral *APC* (do inglês, *Adenomatous Polyposis Coli*), o que leva seus portadores a desenvolverem muitos pólipos adenomatosos em idade precoce (MA *et al.*, 2018).

Conforme ilustrado na figura 1.2, existem três possibilidades de manifestação do CCR: esporádica, familiar e hereditária. A maior porcentagem dos casos se enquadra na forma esporádica, já que ocorre em pacientes sem histórico familiar ou presença de mutações germinativas que predisponham à doença. Por sua vez, a forma familiar é caracterizada por maior incidência da doença em uma determinada família do que o esperado para a maioria da população, embora os indivíduos acometidos não sejam portadores de nenhuma das predisposições genéticas conhecidas. Já a forma hereditária é definida pelo acometimento dos pacientes por anomalias genéticas hereditárias, como a FAP e a Síndrome de Lynch (GIGLIA; CHU, 2016).

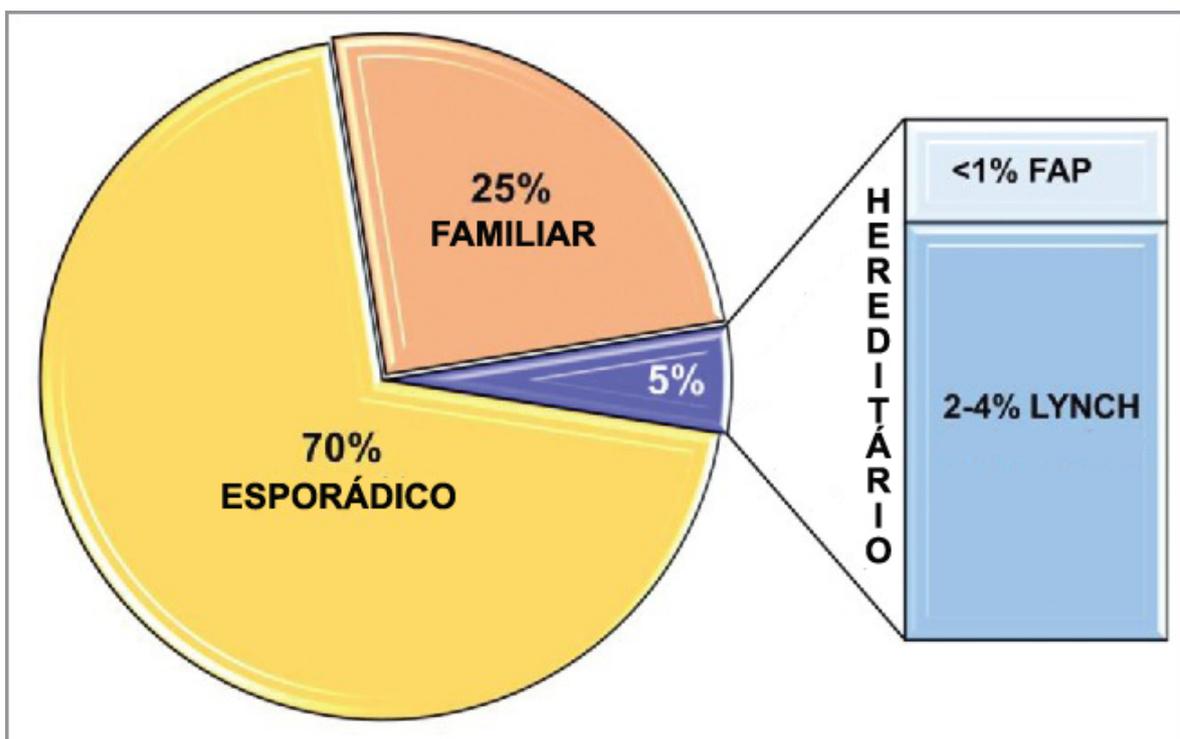


Figura 1.2: Distribuição do CCR de acordo com as manifestações esporádica, familiar e hereditária. O CCR esporádico responde por 70% dos casos, o familiar por 25% e o hereditário por 5%, sendo que em 2-4% dos casos os pacientes são acometidos pela Síndrome de Lynch e em menos de 1% dos casos pela FAP (do inglês Familiar Adenomatous Polyposis). Adaptado de MOKARRAM *et al.*, 2017.

Tendo em vista o exposto, é possível perceber o caráter multifatorial do CCR, o qual sofre influência de fatores ambientais e genéticos. No entanto, para um melhor entendimento dos diferentes aspectos que culminam na aquisição do fenótipo maligno em CCR, é necessário esclarecer alguns mecanismos que constituem as bases celulares e moleculares desta doença.

1.1.2. Bases moleculares e heterogeneidade tumoral

O CCR se desenvolve lentamente ao longo dos anos e é resultado do acúmulo de alterações genéticas e epigenéticas. Um aspecto crucial envolvido na progressão da doença é a instabilidade genômica, a qual pode ser explicada por três mecanismos distintos: instabilidade cromossômica (CIN, do inglês *Chromosomal Instability*), instabilidade de microssatélites (MSI, do inglês *Microsatellite Instability*) e instabilidade epigenética, decorrente do fenótipo metilador de ilhas CpG (CIMP, do inglês *CpG Island Methylator Phenotype*). Em torno de 70% dos casos de CCR de cunho esporádico há CIN, a qual pode ser percebida pelas diferentes alterações que ocorrem em nível cromossômico. Tais alterações podem envolver o cromossomo

inteiro, resultando em aneuploidia (alteração do número de cromossomos), ou podem comprometer a estrutura deles, seja em menor grau devido a mutações que frequentemente resultam em perda de heterozigosidade em genes supressores tumorais, ou em maior grau com a presença de translocações cromossômicas (PINO; CHUNG, 2010; SANSREGRET; VANHAESEBROECK; SWANTON, 2018). Em circunstâncias que se verificam anomalias no mecanismo de reparo de DNA do tipo MMR, regiões de microssatélites são diretamente afetadas, uma vez que os erros que ali ocorrem no decorrer do processo de replicação se acumulam gerando MSI. Conforme descrito anteriormente, a MSI é uma característica da Síndrome de Lynch, no entanto, pode estar presente também em 15% das manifestações esporádicas de CCR. Neste caso, é fruto da inibição da expressão de *MLH1*, devido à hipermetilação da região promotora deste gene, situação que evidencia a associação entre MSI e CIMP (DE' ANGELIS *et al.*, 2018). Finalmente, o CIMP é caracterizado pela presença da hipermetilação global de ilhas CpG em regiões promotoras, resultando na inativação de genes. Além da metilação do promotor de *MLH1* também é comum observar mutação no oncogene *BRAF* (do inglês *v-raf murine sarcoma viral oncogene homolog B*) em amostras consideradas CIMP-high (RHEE; KIM; KANG, 2017).

Como consequência desses processos, genes supressores de tumor sofrem perda de função, enquanto o contrário ocorre com os oncogenes, levando assim a um desequilíbrio na homeostase tecidual (DREWES; HOUSSEAU; SEARS, 2016). Essa quebra na homeostase é percebida por meio de alterações morfológicas e moleculares que ocorrem ao longo do desenvolvimento da doença nas criptas do intestino grosso. A essa sucessão de eventos histopatológicos, primeiramente descritos por Fearon e Vogelstein (1990), deu-se o nome de sequência adenoma-carcinoma. Nesse modelo clássico, a tumorigênese surge a partir de mutações capazes de gerar um estado hiperplásico (com perda do equilíbrio proliferativo sem alterações morfológicas), que é sucedido por um estado displásico (desregulação proliferativa somada a mudanças no fenótipo celular), favorecendo desta forma o surgimento de lesões precursoras inicialmente benignas (pólipos adenomatosos) que evoluem até atingirem o estágio de adenocarcinoma. Inicialmente, mutações no gene supressor tumoral *APC* favorecem o surgimento de adenomas, já que impactam na montagem do complexo responsável pela degradação da β -catenina. Isso permite o acúmulo dessa proteína no citoplasma e sua posterior translocação para o núcleo, onde, ao interagir com os fatores de transcrição TCF/LEF (do inglês

T-cell factor/ lymphoid enhancer binding factor), ativa a transcrição de genes relacionados à proliferação, indiferenciação celular e sobrevivência (HANKEY; FRANKEL; GRODEN, 2018). Também vale ressaltar o papel de APC no estabelecimento da CIN, já que determinadas mutações neste gene podem comprometer a formação do fuso mitótico (ZHANG; SHAY, 2017). Após esse evento, diversas mutações ocorrem em diferentes genes afetando a sinalização mediada por PI3K (do inglês *Phosphoinositide 3-Kinase*), TGF- β (do inglês *Transforming Growth Factor- β*) e receptores tirosina quinase. É interessante evidenciar que mutações acometendo o gene *KRAS* (do inglês *Kirsten rat sarcoma viral oncogene homolog*) e *BRAF* se comportam de forma mutuamente exclusiva, o que evidencia a existência de múltiplas vias de progressão (NETWORK, 2012; GOTO; MARUSAWA; CHIBA, 2013; CARETHERS; JUNG, 2015).

Além da sequência adenoma-carcinoma, outra possibilidade para explicar a carcinogênese colorretal é através da chamada via serrátil. Nesse modelo, a evolução da doença se inicia com o aparecimento de outro tipo de lesão benigna denominada pólipó serrátil. A maioria dos tumores que se desenvolvem através dessa via apresentam mutações em *BRAF* e podem ser associados tanto com MSI quanto com CIMP (DE PALMA *et al.*, 2019).

Como é possível perceber, os mecanismos a partir dos quais os tumores colorretais se desenvolvem podem variar de um paciente para o outro, evidenciando assim a questão da heterogeneidade intertumoral. Por esse motivo, durante mais de uma década, diversos grupos buscaram maneiras de identificar subtipos moleculares do CCR que pudessem guiar a prática clínica, predizendo o desfecho e indicando o tratamento adequado para cada subconjunto de indivíduos. Finalmente, em 2015, um consórcio internacional conseguiu chegar a um consenso e estabeleceu quatro subtipos moleculares, chamados de CMS (do inglês *Consensus Molecular Subtype*). O CMS1 (MSI Imune) apresenta MSI e CIMP-*high*, sofre grande influência do infiltrado imunológico, possui mutações em *BRAF* e é o subtipo com a pior sobrevida após a recorrência. O CMS2 (Canônico) é regido pela via de WNT (do inglês *wingless-type MMTV integration site family*) e apresenta a melhor sobrevida após a recorrência. Já o CMS3 (Metabólico) se caracteriza pela ativação de diversas vias metabólicas e pela presença de mutações em *KRAS*. Por fim, o CMS4 (Mesenquimal) é aquele que apresenta uma assinatura relacionada à EMT (do inglês *Epithelial Mesenchymal Transition*), envolvendo alterações mediadas pela via de TGF- β e presença de infiltrado estromal. Ele configura o subtipo com a pior

sobrevida global e sobrevida livre de recorrência (GUINNEY *et al.*, 2015). A figura 1.3 apresenta um resumo das principais características de cada subtipo.

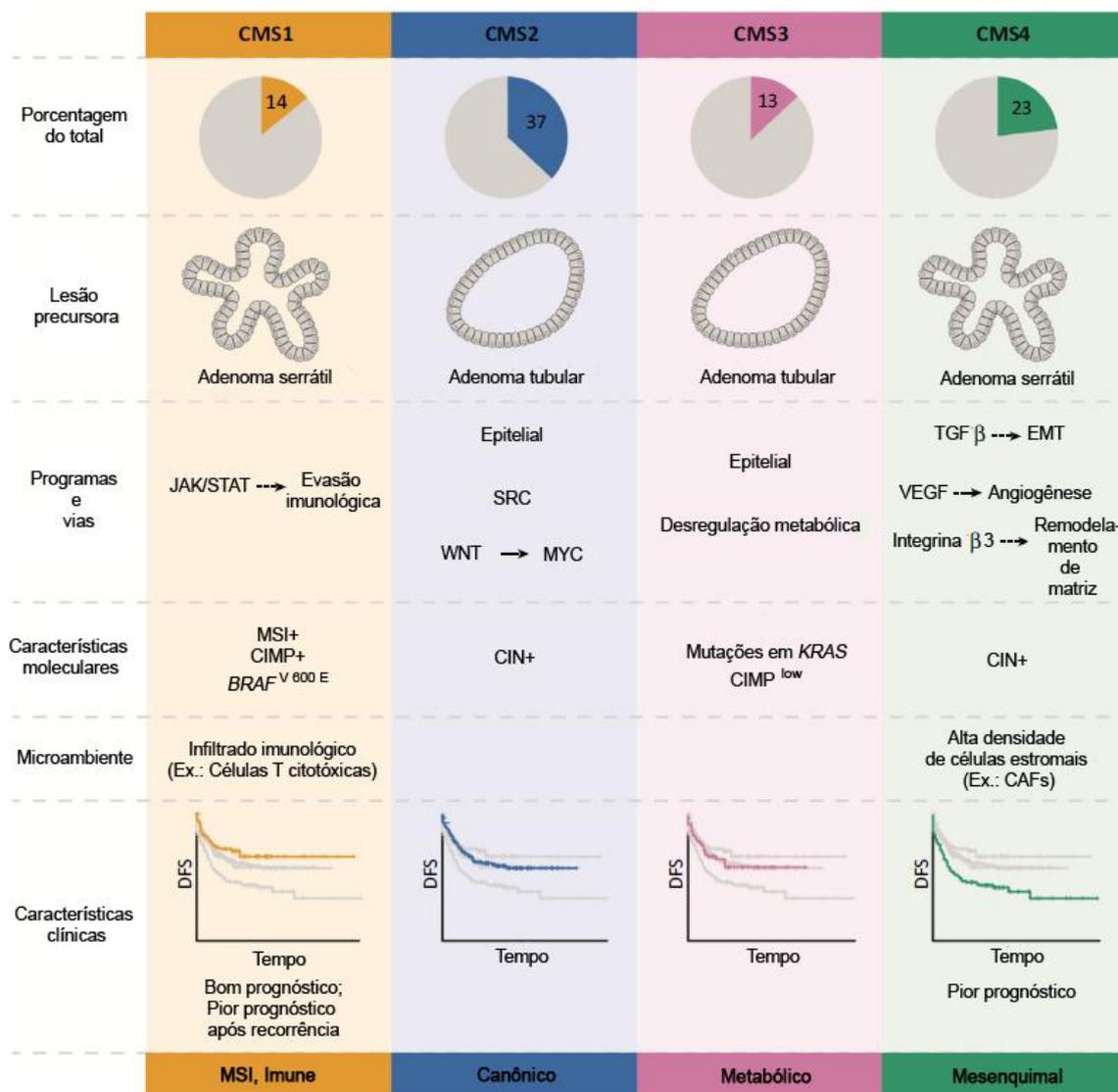


Figura 1.3: Subtipos moleculares do CCR segundo o classificador CMS. A figura mostra os quatro subtipos moleculares de acordo com suas características. Essa classificação leva em consideração aspectos moleculares, clínicos e as principais vias de sinalização acometidas. CAFs (do inglês *Cancer Associated Fibroblasts*); JAK/STAT (do inglês *Janus Kinase/Signal Transducers and Activators of Transcription*); MSI (do inglês *Microsatellite Instability*); CIMP (do inglês *CpG Island Methylator Phenotype*); BRAF (do inglês *v-raf murine sarcoma viral oncogene homolog B*); DFS (do inglês *Disease Free Survival*); SRC (do inglês *v-src avian sarcoma viral oncogene*); WNT (do inglês *wingless-type MMTV integration site family*); CIN (do inglês *Chromosomal Instability*); KRAS (do inglês *Kirsten rat sarcoma viral oncogene homolog*); TGF-β (do inglês *Transforming Growth Factor-β*); EMT (do inglês *Epithelial Mesenchymal Transition*); VEGF (do inglês *Vascular Endothelial Growth Factor*). Adaptado de FESSLER; MEDEMA, 2016.

Embora esse classificador seja capaz de estratificar a maioria das amostras, 13% delas não foram passíveis de classificação o que, em alguns casos, se deu devido à presença de parâmetros comuns a mais de um CMS (GUINNEY *et al.*, 2015). Uma possível explicação para isso poderia estar relacionada à

heterogeneidade intratumoral. Um estudo evidenciou o impacto dessa heterogeneidade nos sistemas de classificação por subtipos ao demonstrar que amostras provenientes de regiões tumorais distintas de uma mesma peça cirúrgica (fronte invasivo e região central), apresentavam um painel de genes diferencialmente expressos capazes de agrupá-las em categorias distintas (DUNNE *et al.*, 2016). Outra limitação do CMS diz respeito ao fato dos subtipos sofrerem influência de fatores oriundos do microambiente tumoral, impossibilitando o estabelecimento de grupos formados somente a partir de características inerentes ao tumor. Com o intuito de solucionar essa questão, pesquisadores utilizando o modelo de PDX (do inglês *Patient-Derived Xenografts*) conseguiram chegar a cinco subtipos denominados CRIS (do inglês *Colorectal Cancer Intrinsic Subtypes*). CRIS-A engloba tumores mucinosos com intensa ativação de vias glicolíticas, presença de MSI ou mutação em *KRAS*. CRIS-B caracteriza-se por tumores com ativação da sinalização por TGF- β e indução de EMT, sendo o subtipo com o pior prognóstico. CRIS-C representa tumores que respondem à inibidores de EGFR, já que são bastante dependentes dessa sinalização. No subtipo CRIS-D, há sinalização mediada por WNT e amplificação, com consequente superexpressão, de *IGF2* (do inglês *Insulin-like Growth Factor 2*). Por fim, CRIS-E tem um fenótipo semelhante às células de Paneth e apresenta mutações em *TP53* (do inglês *Tumor Protein p53*) (ISELLA *et al.*, 2017).

Dados mostraram que o algoritmo de CRIS parece ser superior ao de CMS em relação à classificação de amostras tumorais pareadas (fronte invasivo, região central e metástase linfonodal) de um mesmo paciente (DUNNE *et al.*, 2017) e também em relação à classificação de biópsias parafinadas (ALDERDICE *et al.*, 2018). Além disso, o sistema CRIS mantém a fidelidade de classificação mesmo após o tratamento dos pacientes por três semanas com bevacizumab (anticorpo monoclonal que tem como alvo o VEGF), demonstrando assim a robustez do algoritmo (ALDERDICE *et al.*, 2018).

O amplo debate acerca dos diferentes sistemas de classificação do CCR demonstra a dimensão da complexa arquitetura que rege essa enfermidade. Características inerentes às células tumorais, somadas a fatores relacionados ao microambiente, proporcionam o estabelecimento de subclones heterogêneos dentro da massa tumoral. Assim, ao longo da progressão da doença, células com diferentes arcabouços genéticos e epigenéticos sofrem pressões seletivas oriundas de um microambiente dinâmico, em que há oscilações de nutrientes e oxigênio. A

consequência é a seleção de clones mais adaptados, gerando assim – seja temporal ou espacialmente – a chamada heterogeneidade intratumoral, conforme ilustra a figura 1.4 (MCGRANAHAN; SWANTON, 2017).

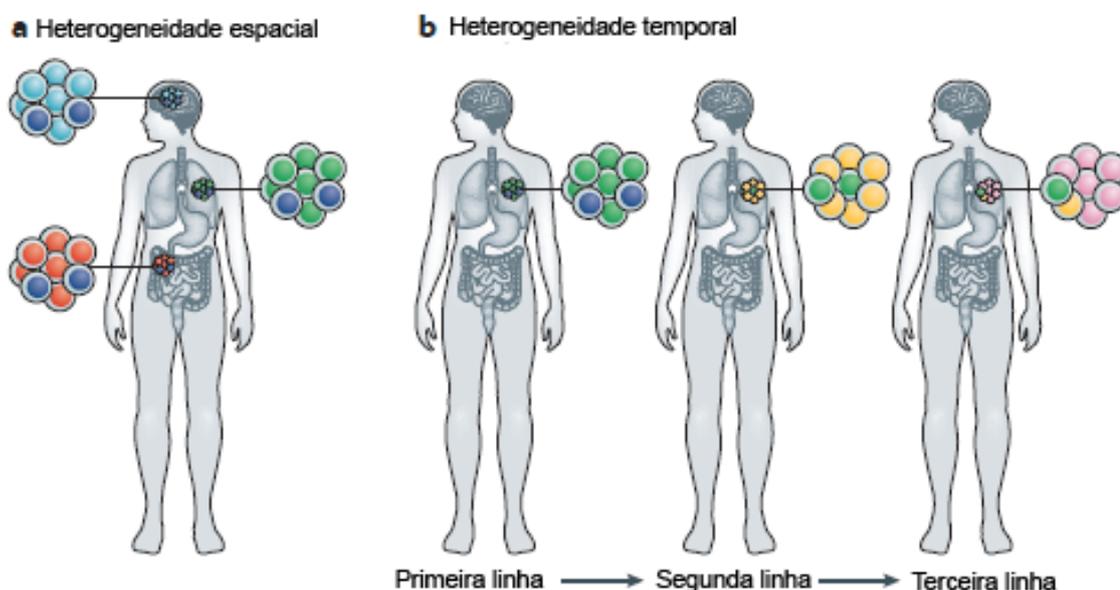


Figura 1.4: Tipos de heterogeneidade intratumoral. A heterogeneidade espacial (a) diz respeito à diferença entre subclones encontrada ao se comparar sítio primário e secundário ou mesmo ao se analisar diferentes regiões da mesma massa tumoral. Já a heterogeneidade temporal (b) pode se referir às mudanças moleculares que ocorrem no tumor ao longo do tempo, seja em relação à evolução da doença ou mesmo em relação à diferentes tratamentos que, conseqüentemente, geram diferentes pressões seletivas. Adaptado de DAGOGO-JACK; SHAW, 2018.

Temporalmente, conforme a doença evolui, um subclone dentro do tumor primário que possua menor capacidade de proliferação pode representar o grupo mais bem adaptado para os processos de disseminação e colonização do sítio secundário e, com isso, ser aquele capaz de dar origem à metástase. De outra maneira, durante um determinado tratamento, pode existir a seleção de subclones que não são afetados pela terapia e, posteriormente, recompõem o tumor a partir de subclones diferentes a cada nova intervenção (DAGOGO-JACK; SHAW, 2018). Por outro lado, espacialmente, a própria composição, seja do tumor primário ou de cada foco metastático, é estabelecida pela presença de clones variados com diferentes capacidades funcionais (KELLER; PANTEL, 2019). Análises de amostras oriundas de diferentes regiões do tumor primário de CCR evidenciaram a existência de diversos clones e subclones dentro da massa tumoral (SOTTORIVA *et al.*, 2015; UCHI *et al.*, 2016; SUZUKI *et al.*, 2017). Além disso, já foi mostrado que metástases linfonodais e distantes podem ser originadas a partir de diferentes subclones do tumor primário (NAXEROVA *et al.*, 2017). Análises do perfil genômico de amostras

pareadas de tumor primário e de metástase pela técnica de WGS (do inglês *Whole Genome Sequencing*) demonstraram que, embora exista uma sobreposição de mutações em ambas as amostras, 15% delas são únicas do tumor primário, enquanto 19% são únicas das metástases e, nesse último caso, parecem impactar vias de sinalização que favoreceriam a colonização de determinado sítio secundário (ISHAQUE *et al.*, 2018).

Em contraste com a ideia de evolução clonal, a heterogeneidade intratumoral também vem sendo discutida sob a ótica das células tronco tumorais (CTTs). Nesta perspectiva, nem todas as células dentro da massa tumoral seriam capazes de dar origem ao tumor, sendo esse papel exercido somente pelas CTTs. Três estudos distintos possibilitaram essa conclusão ao rastrear células *in vivo* através de proteínas fluorescentes. Utilizando um modelo murino sofisticado que combina três aparatos distintos para observar *in situ* o crescimento de adenomas (geração de adenoma intestinal pelo *knockout* de *APC*; identificação por fluorescência do marcador LGR5 de células tronco da cripta intestinal; e sistema R26R-*Confetti* para rastrear as progênes através da mudança de cores induzida pelo tamoxifeno), foi possível observar nos adenomas a mesma hierarquia do intestino sadio com um pequeno grupo de células LGR5+ dando origem a células iguais a elas, mas também a células LGR5- (SCHEPERS *et al.*, 2012). De forma semelhante, outro estudo mostrou que essa hierarquia foi observada em tumores de pele benignos, no entanto, parece ocorrer um deslocamento desse equilíbrio durante a transformação maligna, já que, no carcinoma de células escamosas da pele há um aumento no número de CTTs sem que isso represente um aumento nas células destinadas à diferenciação. Assim, os autores sugerem que a heterogeneidade intratumoral seria decorrente da evolução neutra, uma vez que as CTTs estão envolvidas nos processos aleatórios de divisão simétrica ou assimétrica que ditam o destino das células (DRIESENS *et al.*, 2012). Ou seja, essa heterogeneidade seria resultado desses processos supostamente aleatórios, sem que isso envolvesse qualquer mecanismo de seleção natural. Já o terceiro trabalho mostrou que após a administração de uma droga que eliminava as células proliferativas presentes em gliomas murinos, o crescimento tumoral era retomado por um pequeno grupo de células que expressavam um marcador de células tronco neurais (*nestina*), e davam origem a uma progênie que gradativamente perdia a expressão deste marcador. Entretanto, a eliminação conjunta dessa subpopulação e de suas células progenitoras proliferativas proporcionou a, praticamente completa, eliminação do

tumor (CHEN *et al.*, 2012). Um quarto estudo confirmou que o crescimento de glioblastomas seguia uma hierarquia regida por evolução neutra. No entanto, este mesmo estudo demonstrou a presença de clones que escapavam a essa regra, pois podiam ser expandidos mediante o tratamento com quimioterápicos (LAN *et al.*, 2017).

É interessante ressaltar, no entanto, que essa hierarquia pode ser plástica, pois em determinadas circunstâncias células diferenciadas podem sofrer desdiferenciação e adquirir características de CTTs. Um elegante estudo corroborou para essa conclusão ao documentar que após a depleção de células positivas para o marcador LGR5, em tumores primários de cólon, esta subpopulação era reestabelecida a partir da população negativa para este marcador, o que permitiu novo crescimento tumoral. Por outro lado, o mesmo tipo de depleção em células de tumor secundário (metástase hepática) resultava na sua redução (DE SOUSA E MELO *et al.*, 2017). As diferentes respostas encontradas durante a eliminação das CTTs nos sítios primário e secundário dão indícios de que o microambiente pode desempenhar um papel na ativação de programas que induzem a desdiferenciação. Um exemplo disso foi descrito em ambiente inflamatório onde há presença de TNF- α (do inglês *Tumor Necrosis Factor α*). O TNF- α causa ativação de NF- κ B (do inglês *Nuclear Factor-Kappa B*) e, em contextos em que a β -catenina se encontra constitutivamente ativa, esse fator transcricional causa uma superestimulação da via de WNT que faz com que células não-tronco do intestino apresentem características de CTTs (SCHWITALLA *et al.*, 2013).

Considerando a relevância clínica da heterogeneidade tumoral no âmbito da falha terapêutica e da progressão da doença, é possível presumir que uma melhor compreensão dos mecanismos envolvidos nesse processo pode contribuir efetivamente para o avanço da medicina de precisão.

1.2. Células tronco tumorais

O conceito de CTTs surgiu com o trabalho pioneiro de Bonnet e Dick (1997) em leucemia mieloide aguda, que descreveu nesta doença uma população de células – isoladas a partir do perfil de marcadores de superfície relacionados às células tronco hematopoiéticas (CD34⁺/CD38⁻) – capazes de se diferenciar ao mesmo tempo que mantinham capacidade de autorrenovação quando transplantadas em camundongos imunodeficientes. Essas observações permitiram

afirmar que a leucemia mieloide aguda se estabelecia segundo uma hierarquia. Posteriormente, os mesmos achados foram descritos em câncer de mama, em que o transplante de somente 100 células CD44⁺/CD24^{-/low} em camundongos imunodeficientes era suficiente para iniciar o crescimento tumoral (AL-HAJJ *et al.*, 2003). Essa população também foi descrita em outros tumores sólidos, como câncer colorretal, de ovário, de próstata, de pulmão, dentre outros (BATLLE; CLEVERS, 2017). No entanto, a frequência das CTTs, embora geralmente muito baixa, parece variar de acordo com o tipo de tumor e mesmo dentro do mesmo tipo tumoral. Por exemplo, em adenocarcinoma pancreático essa frequência variou entre 1/2500 e 1/18000 células, já em adenocarcinoma de não pequenas células de pulmão a frequência foi de 1/16000 a 1/30000 células, mas aumentou após sucessivos xenotransplantes (ISHIZAWA *et al.*, 2010).

Como já foi mencionado anteriormente, as CTTs representam uma população de células raras que normalmente são quiescentes e detêm capacidade de autorrenovação, além de darem origem também a células progenitoras, as quais apresentam capacidade transiente de proliferação. Dessa forma, pode-se afirmar que as CTTs obedecem a uma hierarquia (Figura 1.5a) e são responsáveis pelo crescimento em longo prazo do tumor e, por esse motivo, são aquelas que quando xenotransplantadas conseguem promover o crescimento da nova massa tumoral (BATLLE; CLEVERS, 2017).

Baseado em estudos com células tronco adultas de tecidos normais, foi possível inferir que as células tronco podem apresentar dois tipos de divisão celular (Figura 1.5b). Na divisão simétrica, uma célula tronco dá origem a duas células tronco ou duas células progenitoras de amplificação transiente. Já na divisão assimétrica, uma célula tronco dá origem a duas células distintas: uma progenitora e uma tronco. Desse modo, o nicho tem papel fundamental, já que atua ditando qual caminho será seguido por essas células e, conseqüentemente, acaba por promover, uma competição neutra, ou seja, nesta situação não existe célula mais adaptada, mas sim um ambiente com capacidade restrita que desloca o equilíbrio hierárquico mais para um lado ou para outro. Dessa forma, a permanência de uma célula tronco em um determinado nicho é o que garante a ela perpetuação de suas características fenotípicas. No entanto, o que se verifica no câncer (Figura 1.5c) é que as CTTs passam a não mais necessitar dos sinais vindos do nicho, e isso desequilibra a hierarquia celular na massa tumoral, que passa a contar com maior número de CTTs

e menor número de células diferenciadas do que o que se observa em tecidos saudáveis (BATLLE; CLEVERS, 2017).

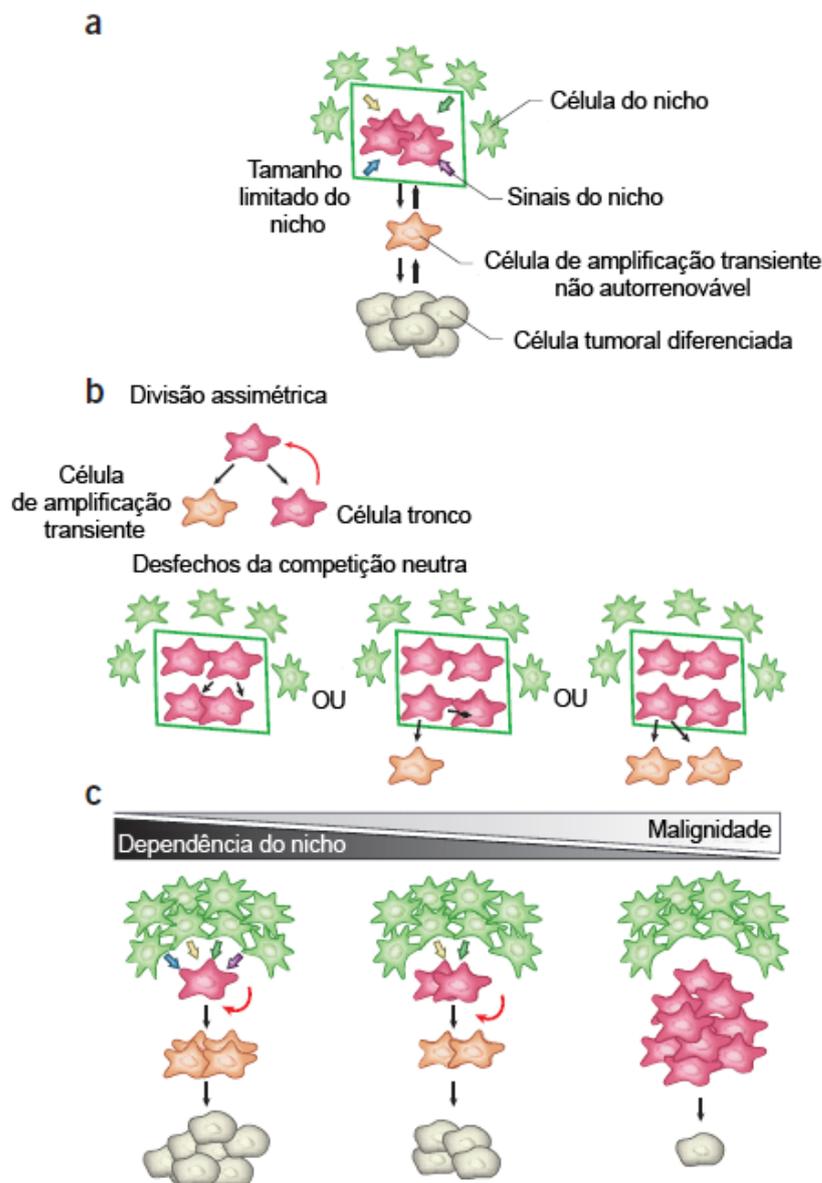


Figura 1.5: Modelo de funcionamento das células tronco em condições normais e tumorais. a) Organização hierárquica das células tronco. b) Maneiras de divisão da célula tronco (simétrica e assimétrica) e influência da competição neutra induzida pelo nicho no destino celular. c) Independência da sinalização do nicho conforme as células tronco adquirem um caráter maligno causando desequilíbrio na hierarquia tecidual. Adaptado de BATLLE; CLEVERS, 2017.

A manutenção das CTTs é regulada por diversas vias de sinalização que estão associadas com a embriogênese, tais como: WNT, JAK/STAT, Sonic Hedgehog e Notch (MATSUI, 2016). Além disso, células tronco embrionárias e tumorais expressam fatores transcricionais relacionados à pluripotência, como SOX2 (do inglês *Sex-determining region Y-box 2*), OCT4 (do inglês *Octamer-binding transcription factor 4*) e NANOG, que são conhecidos por reprimir a diferenciação celular e favorecer o estado tronco (*stemness*) (LIU; YU; LIU, 2013).

Outra característica interessante relacionada às CTTs diz respeito à capacidade que essas células apresentam de resistir às terapias convencionais. Dentre os processos que contribuem para a resistência, é possível citar: a superexpressão de bombas de efluxo (BLEAU *et al.*, 2009), mecanismos de reparo de dano no DNA mais eficientes (DESAI; WEBB; GERSON, 2014), ou até mesmo o aumento da expressão de genes envolvidos com o metabolismo de espécies reativas de oxigênio, que resultam em menor indução de morte dessas células por radiação ionizante, pois esse tipo de morte é mediada pela produção de radicais livres (DIEHN *et al.*, 2009; KIM; KANG; CHO, 2013). Além disso, a plasticidade dessas células possibilita uma alternância dinâmica entre quiescência e entrada no ciclo celular, prevalecendo a manutenção do estado quiescente durante períodos de exposição a tratamentos. Nesse sentido, foi demonstrado que durante o período de tratamento as células tronco de glioblastoma sofrem reprogramação epigenética para favorecer a via Notch, que é capaz de mediar um estado de quiescência. Entretanto, o mesmo estudo observou que alguns clones no tumor primário já apresentam essa reprogramação mesmo antes da exposição do paciente a qualquer droga (LIAU *et al.*, 2017). Esses achados evidenciam que o melhoramento na detecção das CTTs e na compreensão dos mecanismos que regem o estado tronco facilitariam o desenvolvimento de estratégias terapêuticas voltadas para a eliminação dessa população de células, o que resultaria em redução da resistência e, conseqüentemente, da recorrência da doença.

1.2.1. Mecanismos de detecção e isolamento de células tronco

Atualmente, existem diferentes técnicas que permitem a identificação e isolamento da população de CTTs. Todas se baseiam em características intrínsecas dessas células. O ensaio de *side population* utiliza o corante fluorescente *Hoescht 33342* para identificar as CTTs. Como essa população de células apresenta superexpressão de transportadores ABC (do inglês *ATP-Binding Cassette*), responsáveis pela extrusão de drogas, o corante é rapidamente eliminado da célula e a população de CTTs é detectada por citometria de fluxo como um rastro lateral, deslocado em relação à população principal, devido à menor intensidade de fluorescência. Nesse ensaio, o verapamil é usado como um controle, já que, em sua presença, ocorre inibição dos transportadores ABC e a identificação desse rastro é perdida (SHIMODA; OTA; OKADA, 2018). Já o ensaio de formação de esferas

possibilita o enriquecimento de CTTs por meio do crescimento de células em baixa densidade, em um ambiente de não aderência, sem indutores de diferenciação celular. Nesse sistema, apenas células com capacidade de autorrenovação conseguiriam dar origem a uma nova massa tumoral, nesse caso, os esferoides (AKBARZADEH *et al.*, 2019). A técnica considerada padrão ouro para a detecção de CTTs tem sido a tumorigênese *in vivo*, que consiste no xenotransplante de células tumorais em camundongos imunodeficientes. Para isso, as células são inoculadas seguindo uma diluição limitante. Considerando que apenas CTTs poderiam dar origem ao tumor é possível inferir sua frequência (AKBARZADEH *et al.*, 2019). Atualmente, devido às limitações desse ensaio, ele vem perdendo espaço para as técnicas de *lineage tracing*, que consistem no rastreamento das CTTs mediante a expressão gênica de algum marcador específico concomitantemente a uma recombinase, a qual é capaz de ativar um gene repórter. Assim, é possível o estudo *in situ* das CTTs (BATLLE; CLEVERS, 2017). Além dessas técnicas, a citometria de fluxo e o *cell sorting* utilizando anticorpos contra proteínas de superfície presentes em CTTs podem ser empregados. Essas proteínas funcionam como marcadores dessa população, porém devido a sua grande variabilidade entre os diferentes tipos de tumor, ainda não existe um painel universal de marcadores. Alguns desses marcadores são: CD133 (ou Prominina-1), CD44, CD24, LGR5, dentre outros (AKBARZADEH *et al.*, 2019). Utilizando citometria de fluxo, é possível também avaliar a atividade de ALDH1 (do inglês *Aldehyde dehydrogenase 1*), utilizando o kit comercial Aldefluor®, que contém um substrato no qual ALDH1 atua promovendo sua conversão a um produto fluorescente que se acumula nas células. Como as CTTs apresentam alta atividade de ALDH1 é possível identificá-las pela maior intensidade de fluorescência (MELE; LICCARDO; TIRINO, 2018).

Embora diversas metodologias possam ser empregadas para a detecção de CTTs, todas são estabelecidas a partir de características intrínsecas dessa população, seja a síntese de determinada proteína que funciona como marcador, aumento de atividade de uma enzima ou sua capacidade de autorrenovação e de diferenciação. Portanto, avanços no conhecimento dos mecanismos que envolvem a manutenção do estado tronco podem contribuir para uma melhor caracterização dessas células, podendo também levar ao desenvolvimento de novas maneiras de detecção e vice-versa. Nesse sentido, é possível que a glicosilação aberrante tenha um papel significativo tanto na manutenção do perfil tronco, quanto na identificação de novos marcadores, no entanto, este campo vem sendo pouco explorado.

1.3. Glicosilação de proteínas

A glicosilação é um processo de modificação em que carboidratos são adicionados a proteínas ou lipídios por meio de ligação covalente, dando origem aos chamados glicoconjugados. Os carboidratos podem se organizar formando oligo ou polissacarídeos por meio do estabelecimento de ligações glicosídicas entre unidades mais simples chamadas de monossacarídeos. Os monossacarídeos têm fórmula geral $(CH_2O)_n$, com n variando de três a sete. É possível classificá-los de acordo com o número de carbonos ou grupo funcional (aldeído ou cetona) que apresentam. Uma característica fundamental dos carboidratos é a formação de isômeros que podem ser ópticos ou de função. Além disso, os monossacarídeos podem se apresentar na forma cíclica (resultado do ataque que o carbono anomérico sofre por uma hidroxila da própria molécula) que origina isômeros anoméricos alfa e beta (VARKI *et al.*, 2015). Tendo em vista as características apresentadas, é possível perceber que as ligações glicosídicas podem ocorrer em configurações variadas, o que permite o estabelecimento de uma enorme diversidade de moléculas.

Os glicoconjugados podem ser agrupados em diferentes classes: glicoproteínas, proteoglicanos e glicoesfingolipídios (Figura 1.6). As glicoproteínas podem se apresentar como *N*-glicanos, *O*-glicanos ou proteínas ancoradas na membrana por meio de GPI (do inglês *glycosylphosphatidylinositol*). Os proteoglicanos são formados a partir da ligação de cadeias de glicosaminoglicanos (GAGs) a proteínas, com exceção do ácido hialurônico (HA do inglês *Hyaluronic Acid*), que é encontrado livre e apenas interage com proteínas de forma não covalente. Os GAGs que compõem os proteoglicanos, como o heparan sulfato, queratan sulfato e o sulfato de condroitina, são originados a partir da ordenação de dissacarídeos formados por uma glicosamina e um ácido urônico ou galactose. Por fim, os glicoesfingolipídios são formados a partir da ligação de carboidratos a lipídios, nesse caso, ceramidas (VARKI *et al.*, 2015).

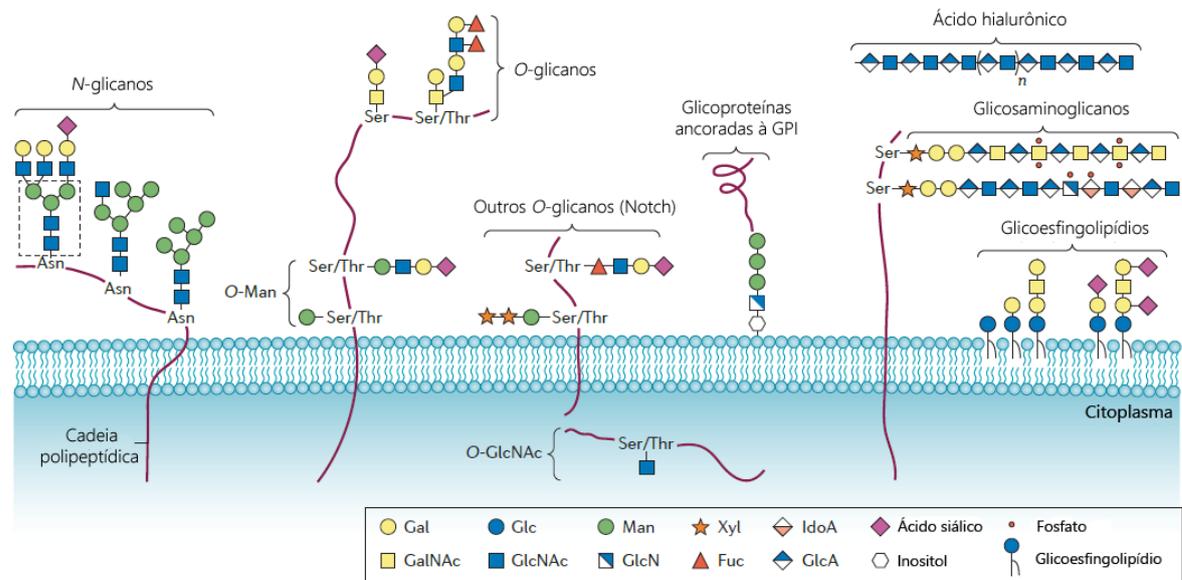


Figura 1.6: Representação das diferentes classes de glicoconjugados presentes nas células. A variedade de glicanos engloba N-glicanos, O-glicanos, glicoproteínas ancoradas por GPI, proteoglicanos formados por glicosaminoglicanos e glicoesfingolípídios. N-glicanos são aqueles ligados a proteínas através da cadeia lateral de resíduos asparagina (Asp) enquanto em O-glicanos essa ligação é feita em resíduos serina ou treonina (Ser/Thr). Os O-glicanos englobam as mucinas, caracterizadas pela presença de uma N-acetilgalactosamina (GalNAc) como primeiro monossacarídeo do glicano e a modificação por O-GlcNAc caracterizada pela ligação dinâmica de uma única N-acetilglicosamina (GlcNAc) em O-glicanos na posição β . Os proteoglicanos são formados por glicosaminoglicanos ancorados a proteínas, com exceção do ácido hialurônico que se encontra livre na matriz extracelular. As proteínas podem ainda ser ancoradas na camada externa da membrana plasmática por meio do glicosilfosfatidilinositol. Por outro lado, a ligação de glicanos a lipídios origina os glicoesfingolípídios. Gal (galactose), Glc (glicose), Man (manose), Xyl (xilose), IdoA (ácido idurônico), GalNAc (N-acetilgalactosamina), GlcNAc (N-acetilglicosamina), GlcN (glicosamina), Fuc (fucose) e ácido glicurônico (GlcA). Adaptado de PINHO; REIS, 2015.

Diferentes reações enzimáticas mediadas por glicosiltransferases e glicosidades, definem o estado de glicosilação das proteínas. Essas enzimas participam, respectivamente, da incorporação e remoção de monossacarídeos à proteínas e glicanos. Para que isso se torne possível, nucleotídeos açúcar devem estar disponíveis para servir de substrato para essas enzimas. Os nucleotídeos açúcar se originam a partir da dieta, da via *de novo* ou da via de reaproveitamento e são resultantes da associação de um monossacarídeo a UDP (uridina difosfato), GDP (guanosina difosfato) ou CMP (citidina monofosfato). Já os monossacarídeos que constituem os nucleotídeos açúcar podem ser: ácido siálico (Neu5Ac), fucose (Fuc), glicose (Glc), manose (Man), galactose (Gal), N-acetilglicosamina (GlcNAc), N-acetilgalactosamina (GalNAc), xilose (Xyl) e ácido glicurônico (GlcA). É importante ressaltar que todos eles podem ser obtidos através da glicose. Findado esse processo, pode-se ter como resultado final a geração de N- e O-glicanos que diferem de acordo com o resíduo de aminoácido da proteína ao qual o monossacarídeo se

liga. Em *N*-glicanos, a ligação ocorre em resíduos de asparagina (Asp) que devem aparecer necessariamente na sequência Asn-X-Ser/Thr (asparagina – qualquer aminoácido – serina ou treonina), enquanto em *O*-glicanos essa ligação ocorre em resíduos serina ou treonina. Outra divergência entre *N*- e *O*-glicanos diz respeito ao seu processo de síntese. Os *N*-glicanos são formados a partir da transferência de um oligossacarídeo precursor (formado por 14 monossacarídeos) em bloco para as proteínas, enquanto em *O*-glicanos os oligossacarídeos são montados pela adição de monossacarídeos um a um (VARKI *et al.*, 2015). Após essa etapa inicial de transferência do oligossacarídeo precursor, os *N*-glicanos são processados e sua maturação envolve diferentes etapas denominadas alongamento, decoração e capeamento. O que se observa então ao final do processo são três tipos principais de *N*-glicanos (Figura 1.7): I) os ricos em manose, que não apresentam antenas formadas a partir da adição de uma molécula de GlcNAc; II) os híbridos, que apresentam uma porção rica em manose e outra na qual ocorre alongamento a partir da adição de uma molécula de GlcNAc, gerando uma antena; e III) os complexos, que se caracterizam pela presença de duas a quatro antenas (VARKI *et al.*, 2015).

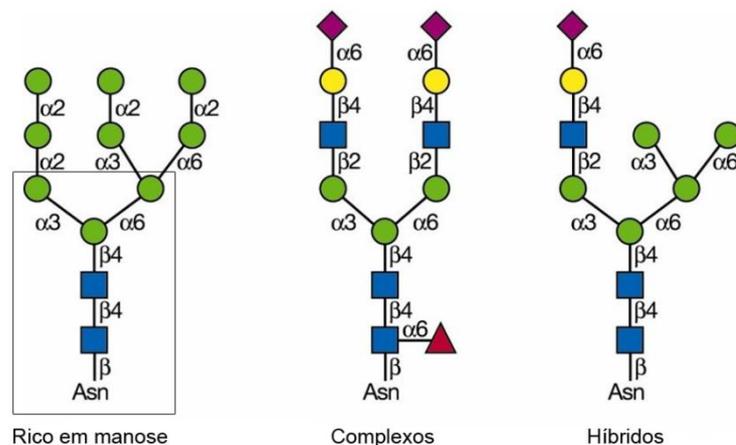


Figura 1.7: Principais tipos de *N*-glicanos. A figura apresenta o *N*-glicanos ricos em manose, híbridos com presença de antena e uma porção de manose e os complexos com suas múltiplas antenas. O retângulo preto ressalta o esqueleto (*core*) invariável dos *N*-glicanos formado pelos 5 monossacarídeos mostrados na figura. Adaptado de VARKI *et al.*, 2015.

Os *O*-glicanos, com exceção de alguns casos especiais, tem sua síntese iniciada a partir da transferência de GalNAc para serinas ou treoninas de proteínas. Frequentemente, a adição de GalNAc é seguida de alongamento desse glicano que pode gerar quatro tipos de estruturas principais, conhecidas como *core*. Assim como ocorre com os *N*-glicanos, esses *core* também podem ser maturados ao longo da sua síntese no complexo de Golgi (BECKER; TRAN; TABAK, 2018).

A modificação por O-GlcNAc representa um tipo especial de O-glicosilação, que envolve a transferência de apenas um monossacarídeo (GlcNAc) para proteínas e pode ocorrer no núcleo, citoplasma ou mitocôndria. A regulação desse tipo de modificação é de responsabilidade de duas enzimas com ações antagônicas: OGT (O-GlcNAc transferase), que catalisa a reação de adição de GlcNAc, e OGA (O-GlcNAcase), que promove a remoção do GlcNAc (Figura 1.8) (NIE; YI, 2019). Para que OGT desempenhe sua função, é necessária a presença de UDP-GlcNAc, nucleotídeo açúcar, que funciona como substrato para a atuação dessa enzima. A síntese do UDP-GlcNAc envolve a integração entre o metabolismo da glicose, de aminoácidos, de ácidos graxos e de nucleotídeos numa mesma via denominada via das hexosaminas (Figura 1.8) e por esse motivo esse nucleotídeo açúcar é considerado um sensor metabólico (HART, 2019).

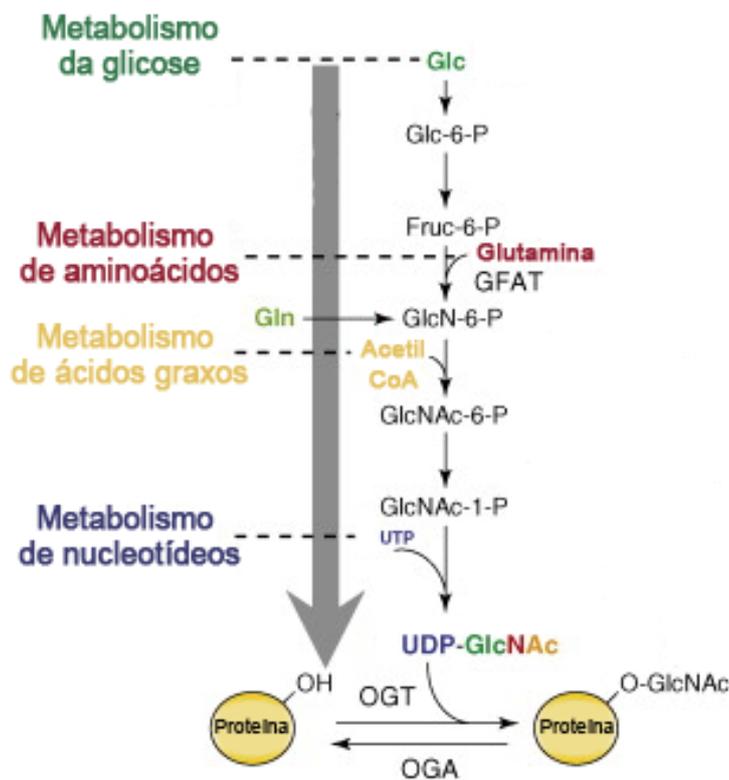


Figura 1.8: Síntese de UDP-GlcNAc através da via das hexosaminas. A figura representa a via das hexosaminas como sensor metabólico, evidenciando a necessidade de influxo de produtos do metabolismo de glicose, aminoácidos, ácidos graxos e nucleotídeos. Essa via depende da enzima limitante GFAT para seu funcionamento e origina como produto final o UDP-GlcNAc. Uma vez sintetizado, o UDP-GlcNAc é utilizado como substrato de OGT para a transferência de GlcNAc para proteínas. Antagonicamente, essa modificação é retirada por OGA. Além disso, o UDP-GlcNAc também é utilizado na síntese de *N*- e *O*- glicanos. Adaptado de SLAWSON; COPELAND; HART, 2010.

1.4. Papel dos glicanos em câncer

Em condições normais, os carboidratos regulam uma série de eventos celulares como: adesão célula-célula e célula-matriz, reconhecimento imunológico de patógenos, sinalização celular promovida por estabilização de proteínas na membrana, dentre outros (CUMMINGS; PIERCE, 2014; CUMMINGS, 2019). Com o estabelecimento de doenças como o câncer, alguns desses processos sofrem desregulação, muitas vezes, mediada por glicosilações aberrantes. A contribuição dos processos de glicosilação ao longo da tumorigênese fez com que esse tipo de modificação pós-traducional fosse apontada como um *hallmark* do câncer (VAJARIA; PATEL, 2017). Nesse sentido, a tumorigênese pode ser favorecida por determinadas estruturas de glicanos, como, por exemplo, é o caso dos O-glicanos truncados, dos antígenos Lewis, dos arranjos sialilados, fucosilados e β -1,6 ramificados.

Durante a síntese de O-glicanos (Figura 1.9b), a adição de GalNAc a resíduos serina/treonina de proteínas origina o antígeno Tn. Após esse evento a enzima C1GALT1, também conhecida como T-sintase, adiciona uma galactose à GalNAc formando o core 1, também conhecido como antígeno T. No entanto, para que isso aconteça, a enzima T-sintase deve apresentar um enovelamento correto que é garantido pela chaperona COSMC. Amostras de CCR apresentam níveis aumentados dos antígenos Tn e da sua forma sialilada sTn (gerada pela ação de ST6GALNAC1) quando comparadas com o tecido normal adjacente, no entanto, não apresentam alterações funcionais nem em COSMC nem em T-sintase (SUN; JU; CUMMINGS, 2018). Além disso, T-sintase também apresenta níveis aumentados em amostras de CCR e foi correlacionada ao pior prognóstico e à aquisição de um caráter maligno pelas células tumorais (HUNG *et al.*, 2014). Somado a isso, há maior atividade de ST3GAL1 e maior marcação do seu produto (sT) em amostras de CCR quando comparadas com o tecido normal adjacente, embora esse efeito seja menos evidente que o observado para os tecidos de câncer de mama (PATIL *et al.*, 2014). Outro aspecto importante está vinculado ao fato dos antígenos sT e sTn serem encontrados predominantemente em MUC1 e na variante 6 de CD44 (HOLST; WUHRER; ROMBOUTS, 2015), uma vez que ambos apresentam correlação com o pior prognóstico dos pacientes com CCR (WANG *et al.*, 2017; LI *et al.*, 2019).

Além dos O-glicanos truncados iniciados pela adição de GalNAc, a modificação por O-GlcNAc (Figura 1.8), um tipo especial de O-glicosilação, também se mostra bastante relevante no contexto do câncer. Esse tipo de modificação depende da disponibilidade do produto final da via das hexosaminas: o UDP-

GlcNAc. Um dos pontos limitantes dessa via é determinado pela enzima GFAT (glutamina-frutose-6-fosfato amidotransferase) que apresenta duas isoformas: GFAT1 (codificada pelo gene *GFPT1*) e GFAT2 (codificada pelo gene *GFPT2*). Ambas têm seus níveis aumentados em amostras de CCR (VASCONCELOS-DOS-SANTOS *et al.*, 2017). Em células de carcinoma de pulmão de não pequenas células estimuladas com TNF/TGF- β , para a indução de EMT, foi observado aumento da expressão de *GFPT2* mediada por NF- κ B o que promovia aumento da migração celular (SZYMURA *et al.*, 2019). Também foi relatado que pacientes com câncer pancreático que apresentavam maiores níveis de GFAT1 foram associados a uma menor sobrevida global (YANG *et al.*, 2016). Outros estudos apontaram aumento dos níveis proteicos de OGT e das modificações por O-GlcNAc em amostras de câncer de pulmão e cólon (MI *et al.*, 2011). O aumento das modificações por O-GlcNAc leva a um aumento dos níveis proteicos de β -catenina, bem como, de sua atividade transcricional resultando em aumento de migração em linhagens celulares de fibroblastos (HAROSH-DAVIDOVICH; KHALAILA, 2018).

A enzima *N*-acetilglicosaminiltransferase-V, conhecida por MGAT5 ou GnT-V, é codificada pelo gene *MGAT5*. Essa enzima é responsável pela transferência de uma *N*-acetilglicosamina para resíduos de manose de *N*-glicanos através de uma ligação β -1,6 a qual origina *N*-glicanos ramificados (Figura 1.9a). Em câncer colorretal, o aumento dos níveis dessa enzima está correlacionado com pior prognóstico e com o desenvolvimento de metástases (MURATA *et al.*, 2000). Com a ramificação do *N*-glicano ele pode ser então alongado através de estruturas LacNAc, compostas por sequências de galactose e *N*-acetilglicosamina geradas principalmente pelas enzimas β 3GNT2 (codificada pelo gene *B3GNT2*) e β 4GALT1 (codificada pelo gene *B4GALT1*). Além dos *N*-glicanos tetra antenados, essas estruturas também podem estar presentes em *O*-glicanos e glicolipídios. Uma vez formadas, elas podem ser decoradas com antígenos Lewis (TOGAYACHI *et al.*, 2010; BYDLINSKI *et al.*, 2018) (Figura 1.9c). O encadeamento de dois ou mais desses dissacarídeos origina as estruturas poli-LacNAc (Figura 1.9a) capazes de interação com galectinas (proteínas com domínios do tipo lectina que se ligam a carboidratos). Essa interação é especialmente importante no caso de receptores de membrana que podem ser estabilizados pelas galectinas, o que prolonga a duração da sinalização celular (FERREIRA *et al.*, 2018). A superexpressão de *B3GNT2* em células de câncer colorretal promove aumento da migração *in vitro* e aumento dos níveis proteicos de MMP-2 e -14, no entanto, as análises de amostras de CCR

indicaram redução da sua marcação em relação ao tecido normal adjacente (JIANG *et al.*, 2018).

Tendo em vista os diversos aspectos da tumorigênese que podem ser impactados pela glicosilação aberrante é plausível hipotetizar que ela também tenha papel sobre a manutenção da população de CTTs.

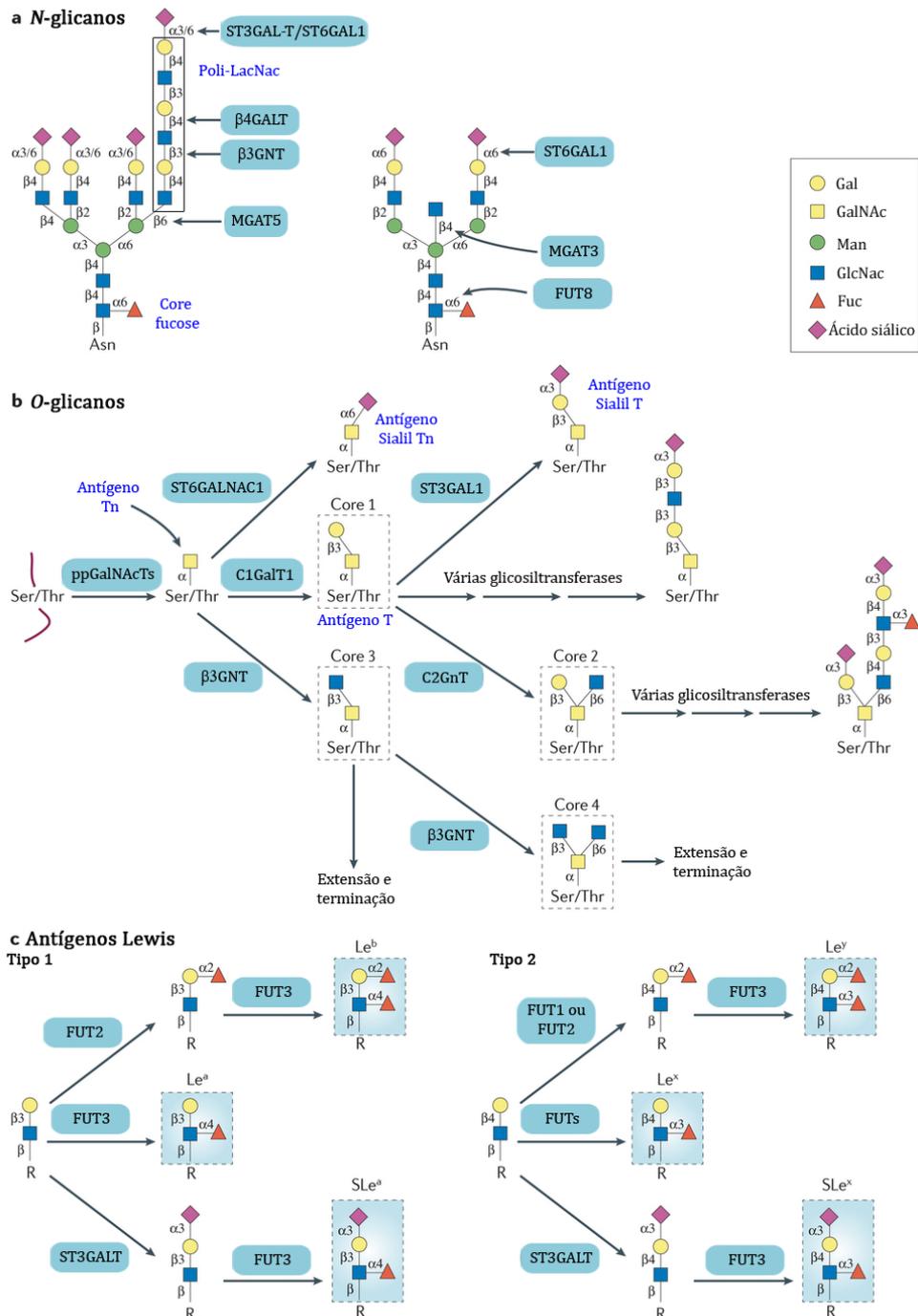


Figura 1.9: Principais estruturas de glicanos relevantes para o câncer. (a) Representa estruturas alteradas em N-glicanos indicando a presença de core fucosilação sintetizada pela enzima FUT8, estruturas Poli-LacNAc sintetizadas pelo trabalho conjunto das enzimas β3GNTs e β4GALTs, sialilação terminal α-2,3 e α-2,6 sintetizadas por ST3GAL4/6 e ST6GAL1, respectivamente e a competição entre MGAT3 e MGAT5 para a síntese de estruturas β-1,4 bifurcadas ou β-1,6

ramificadas. (b) Arranjos de glicanos alterados em O-glicanos destacando a enzima C1GALT1 (também chamada de T-sintase) que necessita da ligação a chaperona COSMC para sua correta função, assim alterações observadas em qualquer uma delas implica em deslocamento da síntese para O-glicanos truncados Tn, sTn. Além desses também merecem destaque o antígeno T e o sT. (c) Estrutura de antígenos Lewis, representando os arranjos de tipo 1 e 2. Caixas verdes representam as glicosiltransferases envolvidas na síntese de cada arranjo. Adaptado de PINHO; REIS, 2015.

1.5. Papel dos glicanos no estado tronco tumoral

Um aspecto interessante que dá indícios que diferentes arranjos de glicanos podem estar envolvidos na regulação da população de CTTs está ligado ao fato da maioria dos marcadores utilizados para identificar essa população serem glicoproteínas (MALLARD; TIRALONGO, 2017).

A modificação de CD44 por O-glicanos truncados em modelos de câncer gástrico aumenta sua interação com o ácido hialurônico da matriz (MEREITER *et al.*, 2019). Por outro lado, a presença de ácidos siálicos terminais em N-glicanos de CD44 impede sua interação com o ácido hialurônico (FALLER; GUVENCH, 2014).

Para a glicoproteína CD133, a perda induzida de um de seus sítios de N-glicosilação impede sua interação com a β -catenina o que resulta em inativação da via de WNT (LIU *et al.*, 2015). Além disso, a presença de sialilação α -2,3, estrutura catalisada por enzimas ST3GAL-T (Figura 1.9a), é capaz de promover redução da cinética de degradação de CD133 (ZHOU *et al.*, 2010).

Para além da atuação dos glicanos em modificações de marcadores de CTTs eles também parecem desempenhar um papel na manutenção da pluripotência. Segundo um estudo realizado com células tronco embrionárias (ESCs, do inglês *Embryonic Stem Cells*) e iPS (do inglês *Induced Pluripotent Stem Cell*), ambas possuíam mais glicanos com sialilação α -2,6 (Figura 1.9a) quando comparadas com células não tronco. A indução de diferenciação reduziu esse tipo de sialilação e, de forma análoga, a remoção dessas estruturas por neuraminidases induziu diferenciação. Vale ressaltar que, além da sialilação, ambas as células também possuíam maiores níveis de fucosilação α -1,3 e α -1,6 (ALISSON-SILVA *et al.*, 2014). Em tecidos normais de cólon, a enzima ST6GAL1 (que promove a sialilação α -2,6) foi documentada apenas na base das criptas (local conhecido por abrigar células tronco), colocalizando com ALDH1. Esse mesmo estudo também demonstrou que células com maiores níveis de ST6GAL1 eram CD133+/ALDH1+ e resistentes ao irinotecano (SWINDALL *et al.*, 2013). De fato, ST6GAL1 parece regular SOX9 para manter o estado tronco (SCHULTZ *et al.*, 2016). Além disso, recentemente, foi

demonstrado que SOX2 é capaz de se ligar ao promotor de *ST6GAL1* e induzir sua transcrição (DORSETT *et al.*, 2019).

A enzima MGAT5 (Figura 1.9a) também parece participar de um mecanismo que beneficia a população de CTTs. Esse mecanismo envolveria a modificação de receptores Frizzled 7 por *N*-glicanos com ramificações do tipo β 1,6-GlcNAc, impactando na sinalização de WNT e resultando no favorecimento da população de CTTs (GUO; NAGY; PIERCE, 2014).

No que diz respeito à fucosilação, um grupo mostrou a partir de um *array* de lectinas feito em células resistentes à gemcitabina que apresentavam características tronco um aumento de fucosilação α -1,2; α -1,3 e α -1,4. Isso foi acompanhado por aumento na expressão de *FUT1*, 2, 3, 4 (Figura 1.9c), além de redução na expressão de *FUT8* (responsável pela adição de fucose na posição α -1,6 ao primeiro GlcNAc do *core* de *N*-glicanos, conforme ilustrado na figura 1.9a). No entanto, o aumento de fucosilação não estava envolvido com a resistência à droga, e foi confirmado ocorrer mesmo em células sem tratamento quando eram submetidas ao ensaio de formação de esferas (TERAO *et al.*, 2015).

Em relação aos O-glicanos, a presença de estruturas truncadas sTn (Figura 1.9b), favorecidas pelo silenciamento da chaperona COSMC, induziu EMT (do inglês *Epithelial Mesenchymal Transition*) em células de câncer pancreático. Estas estruturas foram capazes, também, de promover o fenótipo tronco, o que foi evidenciado pelo aumento de *side population*, de formação de esferas e das células CD44+ ou CD133+, dependendo da linhagem celular estudada (THOMAS *et al.*, 2019). Em CTTs de pâncreas foi observado um aumento de dois antígenos de carboidratos associados ao tumor, Tn e sLe^a (Figura 1.9b e c, respectivamente), e maiores níveis das enzimas GALNT3 (responsável pela adição de GalNAc a resíduos Ser/Thr para síntese do antígeno Tn) e B3GNT3 (responsável pela extensão do core 1 através da síntese da estrutura poli-LacNAc). Além disso, foi observada também redução dos níveis de MGAT4A (responsável pela ramificação β -1,4 que gera *N*-glicanos tri- e tetra-antenados). Curiosamente, tratamentos com inibidores de *N*-glicosilação (como a tunicamicina) e O-glicosilação (como o BAG, benzil-2-acetamido-2-deoxi- α -D-galactopiranosídeo), assim como o silenciamento de GALNT3, foram capazes de promover diminuição dessa subpopulação (BARKEER *et al.*, 2018).

Com relação a modificações por O-GlcNAc (Figura 1.8), foi observado em células de câncer colorretal que a redução na expressão de *OGT* é acompanhada

de redução na população de CTTs as quais passam a expressar *MYBL1*, um ativador transcricional que atua no processo de diferenciação (GUO *et al.*, 2017). Contrariamente, outro estudo, também utilizando células de câncer colorretal, apontou que a inibição de OGT resultou na geração de uma população de células com características tronco que apresentavam aumento dos marcadores CD44+/CD133+ e eram capazes de formar esferas (FUENTES-GARCÍA *et al.*, 2019).

Embora um número discreto de estudos evidencie a presença de glicosilação aberrante em CTTs, nenhum vislumbrou os glicanos como marcadores plausíveis dessa população, tampouco enxergou esses açúcares como sendo fruto de processos alterados que poderiam ressaltar aspectos próprios das CTTs. Além disso, ainda se sabe pouco a respeito do impacto que essas modificações poderiam causar na manutenção do perfil tronco.

2. **Justificativa do estudo**

Globalmente o câncer colorretal representa um importante problema de saúde pública devido a sua alta taxa de mortalidade. A elevada heterogeneidade que perpassa essa doença favorece a recorrência e o desenvolvimento de metástases e é um fator complicador para o estabelecimento de critérios que estratifiquem pacientes por subtipos moleculares capazes de predizer o desfecho e guiar decisões relacionadas à prática clínica.

As CTTs representam uma população com capacidade de autorrenovação e diferenciação capaz de originar todas as células dentro da massa tumoral, além de possuir características intrínsecas que conferem resistência às terapias convencionais. Assim, pesquisas que visem melhorar a identificação das CTTs poderiam indicar grupos de pacientes com maior propensão à recorrência. Tendo em vista que a glicosilação de proteínas é alterada no contexto do câncer, e que muitos dos marcadores utilizados atualmente para a identificação das CTTs são glicoproteínas, a análise mais detalhada dos glicanos diferencialmente expressos em CTTs poderia resultar na identificação de glicobiomarcadores que permitissem discernir entre células tronco adultas normais e tumorais, por meio da identificação de algum carboidrato próprio dessas últimas. Por outro lado, a glicosilação aberrante poderia ter impacto nos mecanismos que sustentam o caráter tronco dessas subpopulações, de modo que uma melhor compreensão desses processos, a partir do ponto de vista das modificações pós traducionais, poderia contribuir para o desenvolvimento de novas estratégias terapêuticas que visem atingir especificamente as CTTs, prevenindo, assim, o desenvolvimento de metástases.

3. **Objetivo**

3.1. Objetivo Geral

Investigar glicobiomarcadores para subpopulações de células tronco tumorais de câncer colorretal.

3.2. Objetivos Específicos

-Estabelecer um protocolo de formação de esferas utilizando células de câncer colorretal.

-Determinar se o ensaio de formação de esferas desenvolvido é capaz de promover enriquecimento de subpopulações de CTTs

-Identificar glicogenes que se encontram diferencialmente expressos nas CTTs.

-Rastrear, *in silico*, possíveis correlações entre a expressão de marcadores de CTTs e glicogenes.

-Comparar, *in silico*, a expressão dos marcadores de CTTs e dos glicogenes entre amostras normais e tumorais, e, entre os diferentes estadios do câncer colorretal.

-Avaliar, *in silico*, o impacto da alta e da baixa expressão de marcadores de CTTs e de glicogenes sobre o desfecho clínico.

4. Materiais e métodos

4.1. Cultura de células

As linhagens celulares derivadas de adenocarcinoma colorretal humano Caco-2 (HTB-37) e HT-29 (HTB-38), obtidas originalmente da *American Type Culture Collection* (ATCC), tiveram suas autenticidades testadas através da análise do perfil de STRs (do inglês, *Short Tandem Repeats*) realizada pelo Laboratório de Bioengenharia Tecidual do Instituto Nacional de Metrologia, Qualidade e Tecnologia (Tabelas 4.1 e 4.2). Essas linhagens apresentam diferenças com relação ao grau de diferenciação, além de apresentarem diferentes perfis mutacionais. Caco-2 é bem diferenciada e apresenta mutação em *APC* e *TP53*. Por outro lado, HT-29 é moderadamente diferenciada e apresenta mutação em *APC*, *BRAF*, *PIK3CA* e *TP53* (AHMED *et al.*, 2013). Ambas foram cultivadas em meio DMEM (do inglês *Dulbecco's Modified Eagle Medium*; GIBCO, Thermo Fisher Scientific) rico em glicose, suplementado com 10% de soro fetal bovino (SFB), 1% de penicilina/estreptomicina (Thermo Fisher Scientific) e mantidas em estufa a 37°C com 5% de CO₂ até atingirem a subconfluência.

Tabela 4.1: Número de repetições em cada alelo da amostra de Caco-2 analisada e da referência (ATCC) para cada um dos sítios genéticos avaliados. O percentual de correspondência entre a amostra e a referência foi de 97 %. O marcador D21S11 não foi utilizado para o cálculo de correspondência, pois não está descrito na amostra referência.

<i>Designação da linhagem celular e.g. nome e número de catálogo</i>	<i>Amostra</i>	
	Caco-2	Caco-2 ATCC® HTB-37
D13S317	13,14	11,13,14
D16S539	12,13	12,13
CSF1PO	11,11	11,11
TH01	6,6	6,6
vWA	16,18	16,18
D21S11	30,30	–
D7S820	11,12	11,12
D5S818	12,13	12,13
TPOX	9,11	9,11
Amelogenina	X,X	X,X

Tabela 4.2: Número de repetições em cada alelo da amostra de HT-29 analisada e da referência (ATCC) para cada um dos sítios genéticos avaliados. O percentual de correspondência entre a amostra e a referência foi de 100 %. O marcador D21S11 não foi utilizado para o cálculo de correspondência, pois não está descrito na amostra referência.

<i>Designação da linhagem celular e.g. nome e número de catálogo</i>	<i>Amostra</i>	
	HT-29	HT-29 ATCC® HTB-38
D13S317	11,12	11,12
D16S539	11,12	11,12
CSF1PO	11,12	11,12
TH01	6,9	6,9
vWA	17,19	17,19
D21S11	29,30	-
D7S820	10,10	10,10
D5S818	11,12	11,12
TPOX	8,9	8,9
Amelogenina	X,X	X,X

4.2. Ensaio de formação de esferas

As células previamente cultivadas em condição 2D (parental) foram tripsinizadas, contadas em câmara de Neubauer e ressuspendidas em meios específicos para a seleção de células tronco tumorais os quais não continham SFB. A composição dos diferentes meios testados é descrita no item a seguir. Além disso, com intuito de promover o cultivo celular em suspensão, foi gerada uma condição de não aderência pelo revestimento prévio de placas de cultura de 6 poços com uma solução de agarose 0,6% contendo 1% de penicilina/estreptomicina. Decorrido o tempo de polimerização da agarose, as células foram plaqueadas e deixadas em cultura por 7 dias, quando eram adicionados mais 1,5 mL de meio de seleção em cada um dos poços. Para ambas as linhagens celulares duas concentrações de células foram testadas: 3×10^3 e 4×10^3 células/poço. Ao final de 14 e 26 dias, respectivamente, as esferas de Caco-2 e HT-29 eram coletadas.

4.2.1. Composição dos diferentes meios testados

Para o ensaio de formação de esferas, três diferentes meios foram testados, a fim de determinar aquele que geraria maior rendimento. Foram eles:

Meio 1 (M1) – DMEM/F12 Ham (Sigma Aldrich) suplementado com 2% de B27 (GIBCO, Thermo Fisher Scientific), 0,5 mg/mL de hidrocortisona (Sigma

Aldrich), 20 ng/mL de FGF (Sigma Aldrich), 20 ng/mL de EGF (Sigma Aldrich) e 1% de penicilina/estreptomicina (Thermo Fisher Scientific).

Meio 2 (M2) – DMEM/F12 Ham suplementado com 2% de B27, 20 ng/mL de FGF, 20 ng/mL de EGF e 1% de penicilina/estreptomicina.

Meio 3 (M3) – DMEM/F12 Ham suplementado com 2% de B27, 20 ng/mL de FGF, 20 ng/mL de EGF, 0,4% de BSA (do inglês *Bovine Serum Albumin*, Sigma Aldrich), 5 µg/mL insulina e 1% penicilina/estreptomicina.

4.3. Análise da expressão gênica por RT-qPCR

4.3.1. Extração e quantificação de RNA

Para a extração do RNA total, as células foram colocadas em contato com Trizol® (Invitrogen) e centrifugadas a 12000 g por 10 minutos a 4°C. A seguir, o sobrenadante foi retirado, sendo posteriormente incubado com 200 µL de clorofórmio por 5 minutos em gelo. Em seguida, as amostras foram centrifugadas a 12000 g por 15 minutos a 4°C. Ao final dessa etapa, a fase aquosa era transferida para um novo microtubo e incubada *overnight* a -20°C com 500 µL de isopropanol. Após esse período, ocorria nova centrifugação das amostras a 12000 g por 10 minutos a 4°C, o sobrenadante era retirado, o pellet restante sofria adição de 1 mL de etanol 75 % diluído em água DEPC (dietilpirocarbonato) e centrifugação a 12000 g por 5 minutos a 4°C. O etanol foi então retirado e as amostras deixadas em temperatura ambiente por 10 minutos para a evaporação do etanol remanescente. Finalmente, o RNA era solubilizado em 10-20 µL de água DEPC e a solução aquecida a 60 °C durante 10 minutos para concluir a solubilização e promover evaporação de etanol residual.

Após a extração, o RNA obtido era quantificado com o auxílio do NanoDrop Lite (Thermo Fisher Scientific), equipamento capaz de mensurar através de espectrofotometria a quantidade de ácidos nucleicos presentes na amostra. Dessa forma, a partir de 1 µL de amostra foi possível obter a quantidade de RNA em ng/µL e o grau de pureza da amostra que é dado pela razão A260/A280. Essa razão aponta possíveis contaminações do RNA com proteínas, já que ácidos nucleicos tem absorvância máxima em comprimento de onda de 260 nm e proteínas em comprimento de onda de 280 nm. Assim, amostras com grau de pureza inferior a 1,7 foram descartadas.

4.3.2. Síntese de DNA complementar (c-DNA)

Após a quantificação, 500 ng de RNA eram tratados com DNase (Invitrogen) durante 30 minutos a 65 °C, a fim de evitar a contaminação da amostra com fragmentos de DNA. A eliminação do DNA foi testada por meio da utilização de amostras que não passavam pelo processo de transcrição reversa, mas eram tratadas ou não com DNase de modo que as tratadas não deveriam apresentar amplificação durante a Reação em Cadeia da Polimerase. Uma vez efetuado esse controle, a fita de c-DNA foi sintetizada a partir da fita simples de RNA por meio de uma reação de transcrição reversa empregando o kit ImProm II ® (Promega). Para isso, eram adicionados ao microtubo contendo 500 ng de RNA tratado com DNase 1 µL de primer randômico (Invitrogen) a uma concentração final de 0,5 µg/µL e 1 µL do mix de dNTP (Invitrogen) a uma concentração final de 10 mM, então a amostra era incubada a 65 °C por 5 minutos seguido de resfriamento em gelo por 1 minuto, depois eram adicionados 4 µL do tampão 5x, 2,5 µL de MgCl₂ (25 mM) e 0,5 µL de água livre de nuclease. Era feita uma incubação de 2 minutos a 65 °C. Por fim, foi adicionado 1 µL da enzima transcriptase reversa ImProm II ® e as amostras foram colocadas em termociclador com a seguinte ciclagem: 25 °C - 5 minutos/ 42 °C - 60 minutos/ 70 °C - 15 minutos.

4.3.3. Obtenção das sequências iniciadoras

A sequência dos iniciadores de *LGR5* foi obtida a partir de dados da literatura (LIN *et al.*, 2015) enquanto as demais sequências dos marcadores de CTTs foram gentilmente cedidas pela Dra. Sheila Coelho. As sequências dos glicogenes foram todas desenhadas conforme descrito a seguir. Inicialmente as sequências dos mRNAs de cada glicogene estudado foram obtidas em formato FASTA através da plataforma do NCBI (www.ncbi.nlm.nih.gov/gene). Todas as variantes eram então aneladas por meio da ferramenta *online* Clustal Ω (www.ebi.ac.uk/Tools/msa/clustalo/) e áreas de sobreposição entre elas foram selecionadas. Utilizando a plataforma Ensembl (www.ensembl.org) a sequência dessas áreas era então verificada para confirmar se correspondia a uma região de éxon. Feito isso, os iniciadores eram desenhados pelo *software* Generunner seguindo os seguintes critérios: tamanho de 18 a 25 pares de bases, temperatura de dissociação (*melting*) variando de 56 à 60 °C, presença de 40-60% de citosina e

guanina, temperatura para a formação de *hairpin* de, no máximo, 30 °C e variação da energia livre de Gibbs (ΔG) que mede a espontaneidade da ocorrência de uma reação sendo até -10kcal/mole, tanto quando se considerava a formação de homodímeros ou heterodímeros e, por fim, apresentar um amplicon de 80 à 150 pares de bases. Os parâmetros eram então reavaliados com a plataforma IDT OligoAnalyser (www.idtdna.com/calc/analyzer). Finalmente, a especificidade de cada uma das sequências era verificada *in silico* através da plataforma BLAST (do inglês *Basic Local Alignment Search Tool* < www.blast.ncbi.nlm.nih.gov/Blast.cgi>). As sequências iniciadoras são apresentadas na tabela 4.3.

Tabela 4.3: Lista contendo as sequências iniciadoras utilizadas nas análises de qPCR.

Genes	Sequência (5'-3')
PROM1 (CD133)	Senso: TCATACTGGTGGCTGGGTGG Antissenso: GGTGGTCGGGGTGGCAT
CD44	Senso: CTTCAGGAGGTTACATCTTTTACAC Antissenso: TCATCAAAGTGGTAGCAGGGA
CD24	Senso: TGCTCCTACCCACGCAGATT Antissenso: GGCCAACCCAGAGTTGGAA
LGR5	Senso: GTGGCAGCAAGTATGGCG Antissenso: AGCAAAGGGAATTGAGCAAG
ALDH1	Senso: TGCTGGCGACAATGGAGTCAATG Antissenso: AACCTGCACAGTAGCGCAATGT
B3GNT2	Senso: CGAGGCATACTGGAACCGAGAG Antissenso: CAGGTCAGGTTTCGCAGTAGTTCAG
B4GALT1	Senso: CCGCTAGCAACTTGACCTCG Antissenso: CTCCAGGTCCACAGGCATGTTAAAC
MGAT5	Senso: GGCAGAAAAGCAGAACCTTG Antissenso: AGCTGAAATCCTAATGTCATGGC
GFPT1 (GFAT1)	Senso: TCCGGCATCATGTGTGGTATATTTG Antissenso: GCCTCCATCAAATCCCACACC
GFPT2 (GFAT2)	Senso: TGCCTACATGAACTACAGAGTCC Antissenso: CCGAGTCGTAGCCTCTGTAC
OGA	Senso: GCAGTGCGGATAAACAGGAAGC Antissenso: CAAGGGTCCTGTCTCGTTCTC
OGT	Senso: GCCGCCATTTCAAGACCGTAC Antissenso: CTACCAGCAAGAACGTATTTGGAGG
ST3GAL1	Senso: CCAAGCAGATGGTCCTGGA

	Antissenso: CATCGGTAGGTGTCGTCCTC
PPIA	Senso: AACCACCAGATCATTCTT Antissenso: GCGAGAGCACAAAGATTC
UBE2D2	Senso: TGTTCTCTGTTGTGTGATCCCA Antissenso: CCCGAGCTATTCTGTTGACTTT
ACTB	Senso: TACAATGAGCTGCGTGTGG Antissenso: TAGCACAGCCTGGATAGCAA
18S	Senso: AACCCGTTGAACCCATT Antissenso: CCATCCAATCGGTAGTAGCG
PGK1	Senso: CAAGCTGGACGTTAAAGGGA Antissenso: CTTGGGACAGCAGCCTTAAT
ZUP1 (ZUFSP)	Senso: GCAGTCTCAAGCAACTTCGG Antissenso: TGTCTCCTGGCAAGTTTCTCC
LRRC2	Senso: GGGTGTGGAGGCATCATCT Antissenso: TGCAATTGTTGAAGGCTGGG
B2M	Senso: TCCAGCGTACTCCAAAGATTCA Antissenso: ATGTCGGATGGATGAAACCC
GUSB1	Senso: TCGATGACATCACCGTCACC Antissenso: GTCCATTCGCCACGACTTT
GAPDH	Senso: CAACAGCCTCAAGATCATCAGCAA Antissenso: AGTGATGGCATGGACTGTGGTCAT

4.3.4. Reação de PCR quantitativo em tempo real (qPCR)

Primeiramente, as reações foram feitas empregando diluições seriadas do c-DNA a fim de gerar uma curva padrão para as sequências iniciadoras. A partir daí, foi possível obter uma regressão linear da curva padrão a qual fornece informações a respeito da linearidade (R^2), inclinação da reta (*slope*) e eficiência da reação (dada através da fórmula $E=10^{(-1/slope)} - 1$). Tendo em vista isso, os oligonucleotídeos que apresentaram eficiência de reação entre 90-110 % e coeficiente de determinação (R^2) maior que 98% foram utilizados.

A RT-qPCR foi realizada nos termocicladores *Applied Biosystems* 7500 e ViiA 7 (Thermo Fisher Scientific). Reações com 10 μ L de volume final eram constituídas de: 1 μ L de cDNA diluído 1:2; 0,6 μ L de iniciadores a uma concentração final de 0,3 μ M; 5 μ L do reagente GoTaq® qPCR Master Mix (Promega) que emite fluorescência mediante interação com as moléculas de dupla fita; e 3,4 μ L de água livre de nuclease. Cada reação foi feita em triplicata em presença de dois controles

negativos para cada alvo. As reações envolvendo os marcadores de CTTs tinham o seguinte perfil de ciclagem: 95 °C - 30 segundos/ 60 °C - 40 segundos/ 60 °C - 30 segundos durante 40 ciclos, com exceção de *LGR5* que ciclava à 95 °C - 30 segundos/ 55 °C - 40 segundos/ 72°C - 30 segundos também por 40 ciclos. Já o perfil de ciclagem dos glicogenes era: 95 °C - 30 segundos/ 60 °C - 40 segundos/ 72°C - 30 segundos durante 40 ciclos.

A expressão dos genes alvo foi normalizada pela média geométrica dos valores dos genes referência *PPIA* e *UBE2D2*, no caso das análises referentes à Caco-2, enquanto para as análises referentes à HT-29 a normalização foi feita apenas pela média aritmética dos valores de *PPIA*. O cálculo da expressão gênica foi realizado utilizando a fórmula $2^{-\Delta\Delta CT}$ (LIVAK; SCHMITTGEN, 2001; SCHMITTGEN; LIVAK, 2008).

4.4. Análises in silico

4.4.1. Análise comparativa da expressão gênica entre amostras normais e tumorais

Para este fim foi empregada a plataforma UALCAN (CHANDRASHEKAR *et al.*, 2017) por meio da qual é possível analisar diversos parâmetros utilizando amostras oriundas do banco de dados TCGA (do inglês *The Cancer Genome Atlas*). As amostras normais e tumorais depositadas no banco de dados de adenocarcinoma de cólon (COAD) e de adenocarcinoma de reto (READ) foram analisadas quanto à expressão dos marcadores de CTTs (*CD24*, *CD44*, *LGR5*, *PROM1*, *ALDH1*) e glicogenes (*B3GNT2*, *B4GALT1*, *GFPT1*, *GFPT2*, *MGAT5*, *OGA*, *OGT*, *ST3GAL1*). Também foi avaliada a expressão dos genes que apresentaram alteração comparando amostras provenientes de diferentes estadios do CCR.

4.4.2. Análises de correlação

A verificação de possíveis correlações lineares entre a expressão dos marcadores de CTTs e das glicogenes foi realizada através da plataforma GEPIA (TANG *et al.*, 2017) que, assim como a anterior, também utiliza o banco de dados do TCGA. Dessa forma, as amostras tumorais depositadas nos bancos COAD e READ foram avaliadas por meio de gráficos de correlação que continham em um dos eixos

dados de expressão gênica de um dos marcadores de CTTs e no outro eixo dados, também de expressão gênica, de um dos glicogenes avaliados.

4.4.3. Análises do desfecho clínico

Ainda fazendo uso do GEPIA, os dados dos bancos COAD e READ foram classificados ou não de acordo com seu perfil molecular (MSS ou MSI-*high* e MSI-*low*) e foram divididos em tercils de acordo com a expressão tanto de marcadores de CTTs quanto de glicogenes. Desse modo, foi possível estabelecer grupos com alta expressão, quando pertenciam ao tercil superior, e baixa expressão, quando pertenciam ao tercil inferior. Utilizando esses critérios foi estabelecida então a sobrevida global e a sobrevida livre de doença, ambas em um período de 4000 dias através da curva de Kaplan-Meier.

4.5. Análise estatística

Para a realização das análises estatísticas foi utilizado o *software* GraphPad Prism 5.0. Dados provenientes de sete ou mais experimentos foram submetidos ao teste de normalidade de Shapiro-Wilk. Para observações pareadas que seguissem uma distribuição normal era aplicado o teste t de Student. Para as análises utilizando dados do TCGA foram empregados diferentes testes estatísticos. Quando foram comparadas amostras normais e tumorais foi empregado o teste t de Student. Para as análises de correlação foi determinado o coeficiente de correlação de Pearson o qual é capaz de medir o grau de associação linear entre duas variáveis. Com relação aos dados de sobrevida, foi aplicado o teste *log rank*, a fim de estabelecer uma comparação entre as curvas. Além disso, também se estimou a probabilidade de um indivíduo sofrer o evento (morte ou recorrência da doença no caso de SG ou SLD, respectivamente) a qual é dada pelo valor de HR (do inglês, *Hazard Ratio*) obtido pela regressão de Cox. Os dados foram considerados estatisticamente significativos quando: * $p < 0.05$; ** $p < 0.01$ e *** $p < 0,001$.

5. Resultados

5.1. Estabelecimento do ensaio de formação de esferas para enriquecimento de CTTs.

A fim de estabelecer os parâmetros ideais para o ensaio de formação de esferas, foram feitos diferentes testes em que a linhagem HT-29 de adenocarcinoma de cólon humano era cultivada em placas de cultura com poços de 8,962 cm² previamente revestidos com agarose. A geração de esferas foi monitorada por microscopia de contraste de fase a partir do plaqueamento de duas quantidades de células por poço (3×10^3 e 4×10^3) em três diferentes meios de cultura (composição descrita no item 4.2.1), denominados M1, M2 e M3. As figuras 5.1, 5.2 e 5.3 ilustram o acompanhamento da formação de esferas realizado a partir do plaqueamento de 3×10^3 células/poço em presença dos meios M1, M2 ou M3, respectivamente. Já as figuras 5.4, 5.5 e 5.6 ilustram o acompanhamento da formação de esferas a partir do plaqueamento de 4×10^3 células/poço em presença dos meios M1, M2 ou M3, respectivamente. No que diz respeito à quantidade de células HT-29 plaqueadas, a utilização de 4×10^3 células/poço se mostrou mais efetiva, sendo possível observar a presença de esferas já nos primeiros dias de monitoramento. Dos meios utilizados, o meio M3 parece ser o menos propício para a formação de esferas ao final de 31 dias, enquanto o meio M1 foi o que induziu a formação de esferas maiores, indicando ser ele o mais adequado para o ensaio. É possível perceber também que, fazendo uso dos meios M1 ou M2, houve uma redução do tamanho e da quantidade de esferas observada no dia 10 em relação ao que havia sido observado no dia 7. Uma possível explicação para esta redução seria o esgotamento dos suplementos do meio em torno de 7 dias. Considerando essa hipótese, foi realizado um ensaio de formação de esferas utilizando as células HT-29 (4×10^3 células/poço em presença de meio M1) em que após 7 dias de cultivo houve adição de 1,5 mL do meio M1 (Figura 5.7). Considerando a maior eficiência desta estratégia para a formação de esferas, em relação ao que havia sido observado nos testes anteriores, a mesma foi utilizada nos experimentos subsequentes. É importante ressaltar também que para a linhagem HT-29, ficou estabelecido um tempo de cultivo de 26 dias para a obtenção das esferas, tendo em vista que após esse período foi verificado que as esferas passavam a aderir na agarose que reveste a placa de cultura, o que pode ser visto na imagem do dia 31 (Figura 5.4).

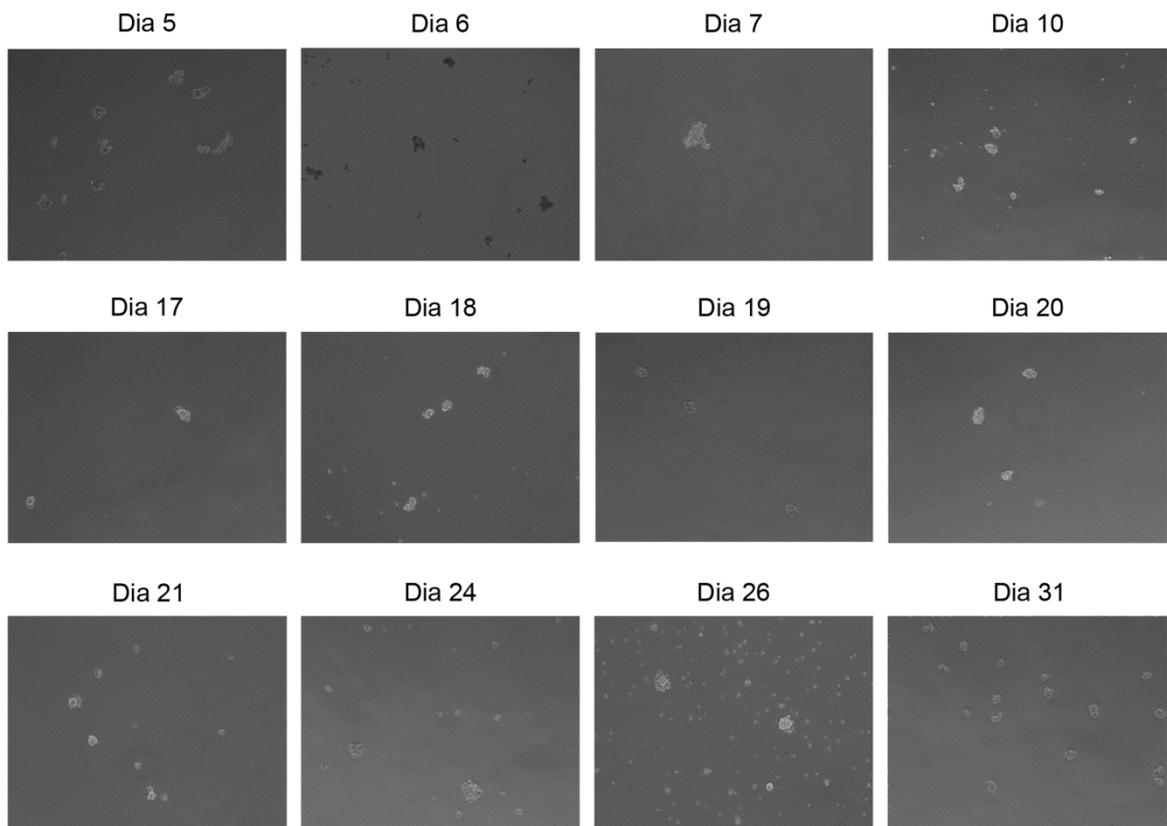


Figura 5.1: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (3×10^3 células/poço) em meio M1. Imagens representativas do monitoramento morfológico por microscopia de contraste de fase durante um período de 31 dias. Aumento: 100X

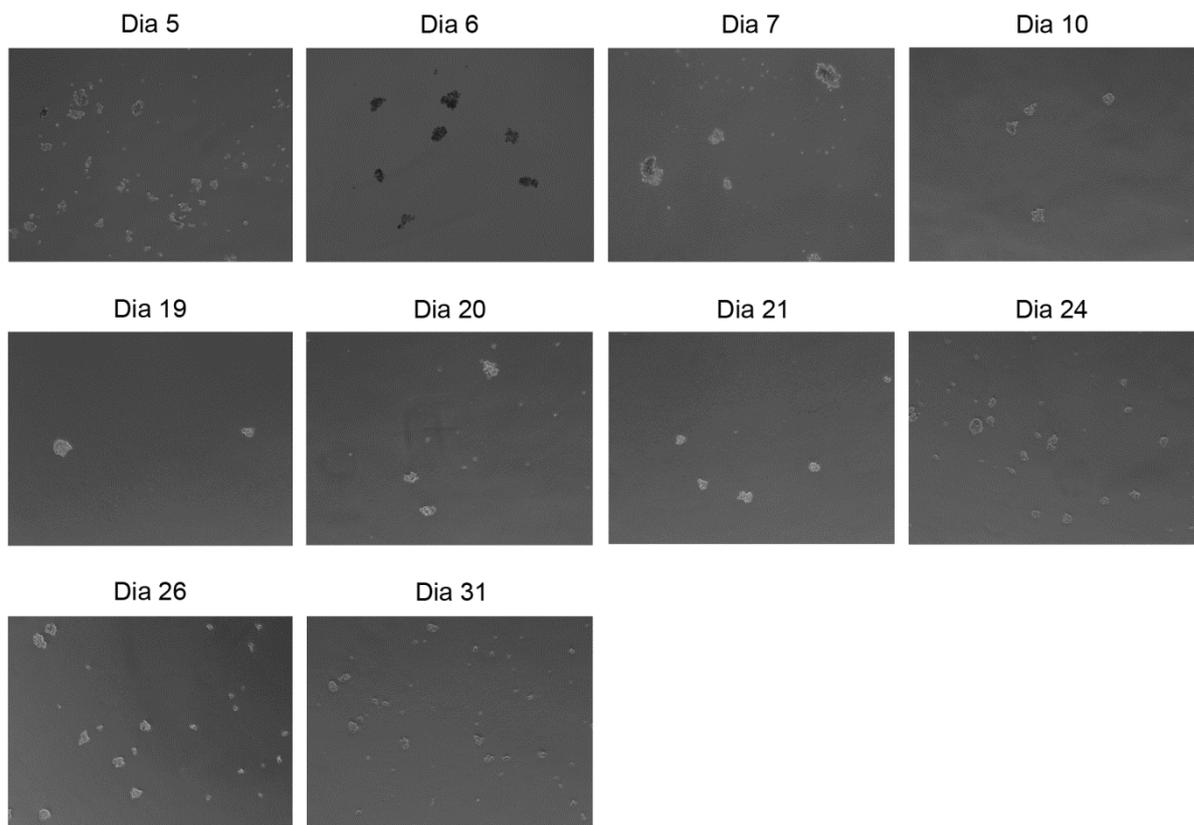


Figura 5.2: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (3×10^3 células/poço) em meio M2. Imagens representativas do monitoramento morfológico por microscopia de contraste de fase durante um período de 31 dias. Aumento: 100X

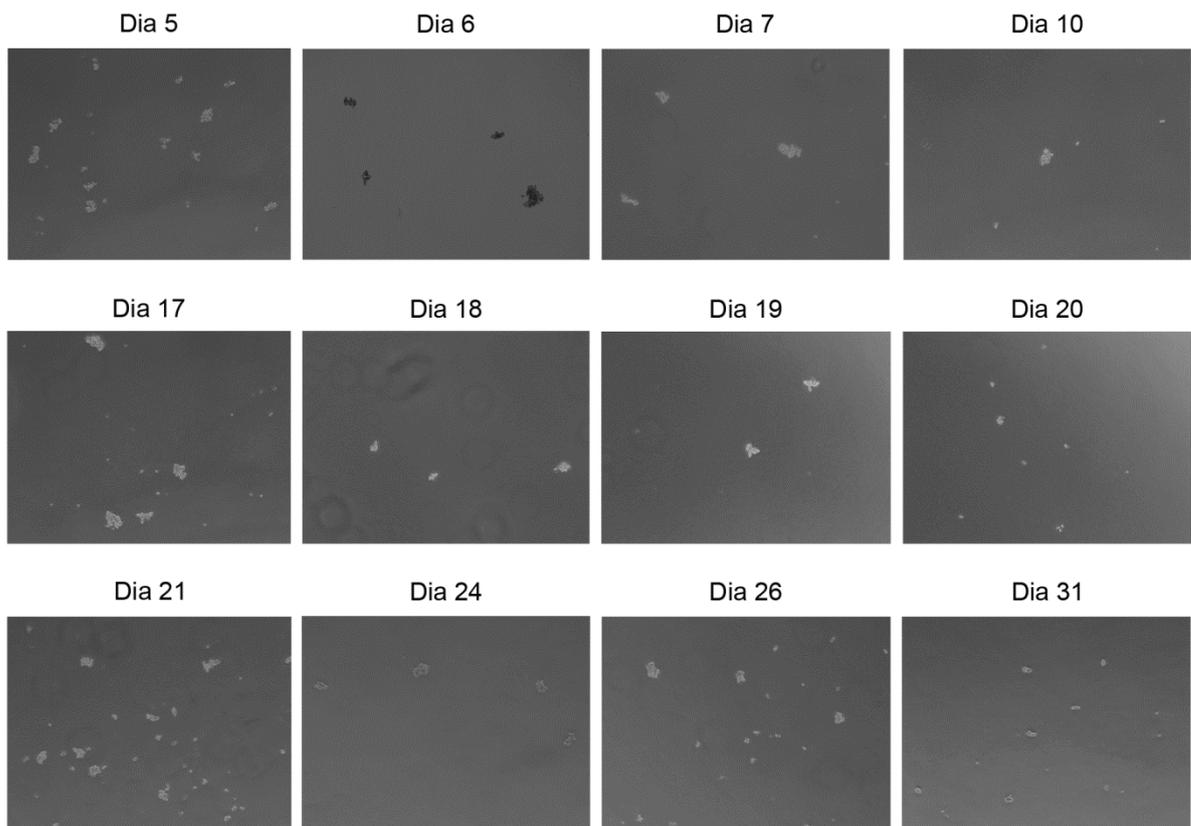


Figura 5.3: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (3×10^3 células/poço) em meio M3. Imagens representativas do monitoramento morfológico por microscopia de contraste de fase durante um período de 31 dias. Aumento: 100X

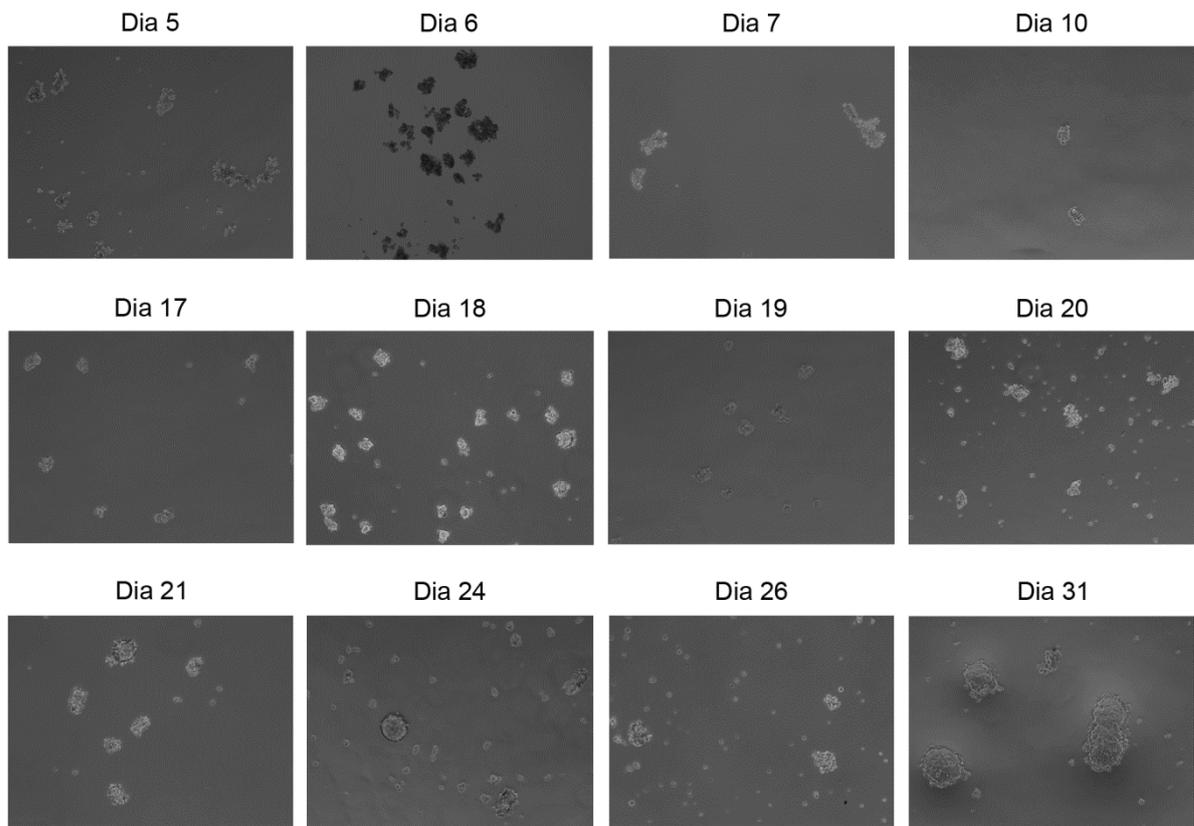


Figura 5.4: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (4×10^3 células/poço) em meio M1. Imagens representativas do monitoramento morfológico por microscopia de contraste de fase durante um período de 31 dias. Aumento: 100X

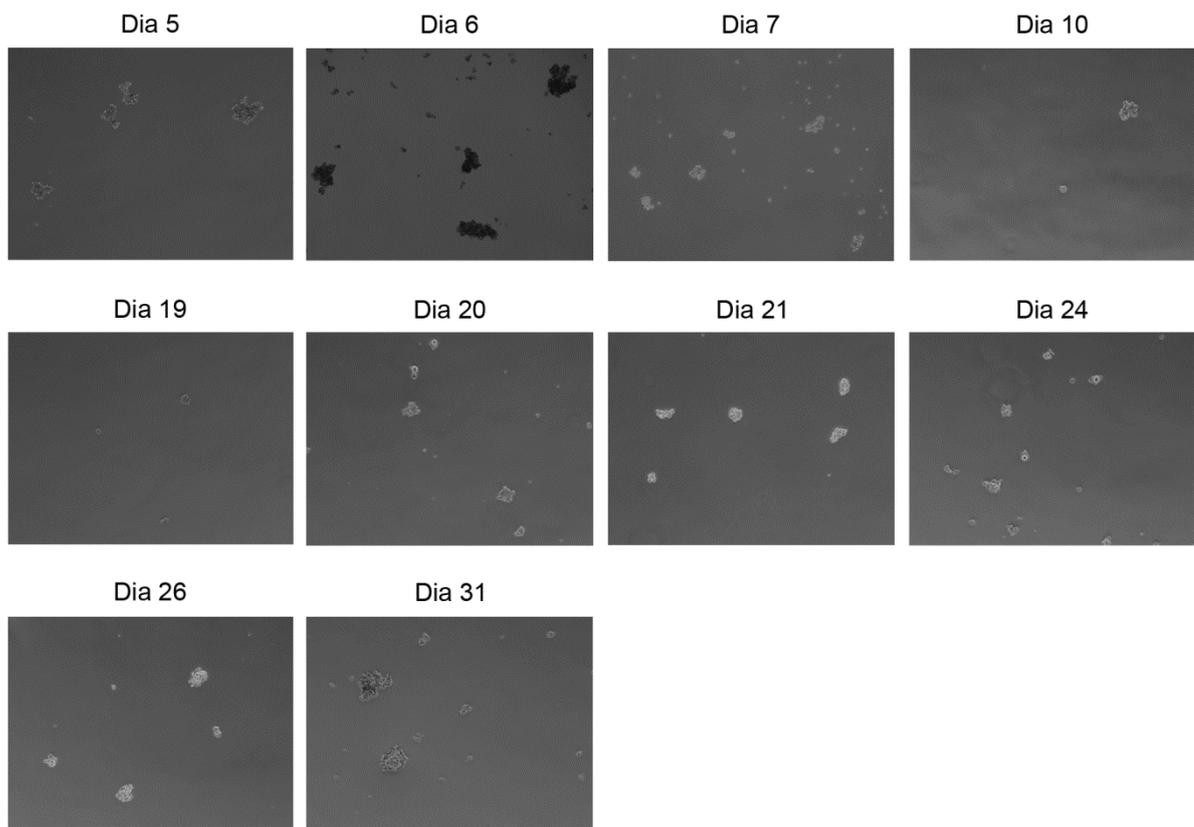


Figura 5.5: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (4×10^3 células/poço) em meio M2. Imagens representativas do monitoramento morfológico por microscopia de contraste de fase durante um período de 31 dias. Aumento: 100X

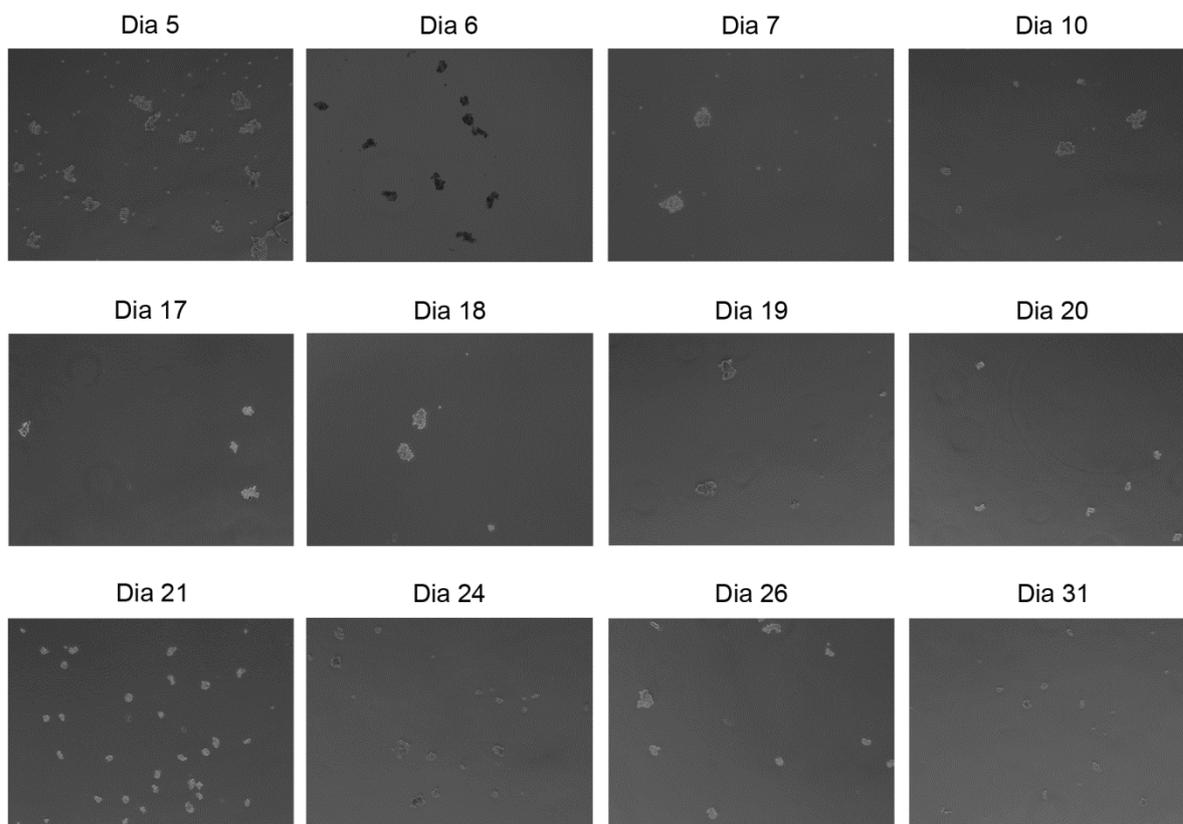


Figura 5.6: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (4×10^3 células/poço) em meio M3. Imagens representativas do monitoramento morfológico por microscopia de contraste de fase durante um período de 31 dias. Aumento: 100X

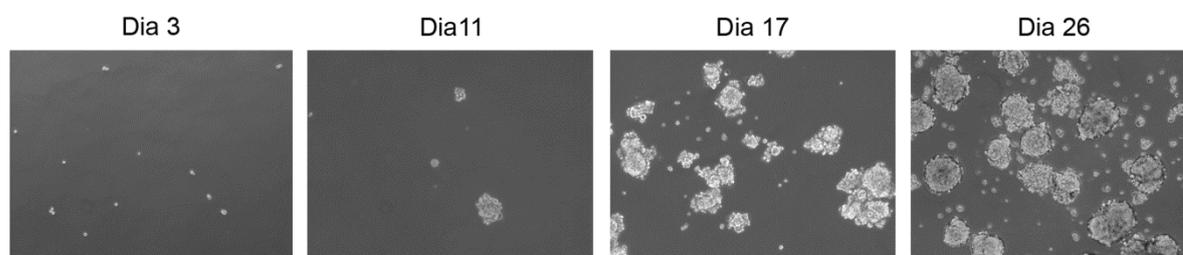


Figura 5.7: Acompanhamento da formação de esferas a partir do plaqueamento das células HT-29 (4×10^3 células/poço) em meio M1 seguindo o protocolo padronizado. Imagens representativas do monitoramento morfológico por microscopia de contraste de fase durante um período de 26 dias. Aumento: 100X

Uma vez estabelecido o protocolo para a linhagem HT-29, o mesmo foi empregado para a linhagem Caco-2 (Figura 5.8). No entanto, Caco-2 forma esferas em menos tempo que HT-29. Neste caso, o tempo ideal de cultivo foi de 14 dias. Após esse período, foi possível observar o aparecimento de regiões mais escuras nas esferas, que chegaram a ocupar toda a esfera no dia 26. Essas regiões representam áreas de necrose e por isso, é imprescindível trabalhar em tempos anteriores ao seu surgimento.

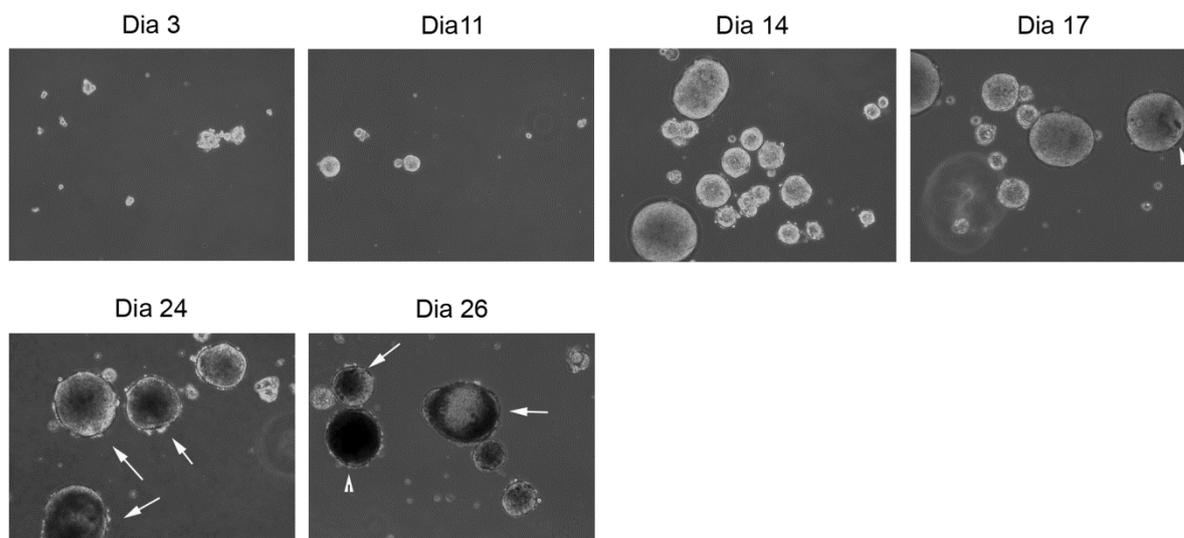


Figura 5.8: Acompanhamento da formação de esferas a partir do plaqueamento das células Caco-2 (4×10^3 células/poço) em meio M1 seguindo o protocolo padronizado. Imagens representativas do monitoramento morfológico por microscopia de contraste de fase durante um período de 26 dias. Setas, esferas com áreas de necrose; cabeça de seta, área de necrose ocupando toda a esfera. Aumento: 100X.

Portanto, o meio mais adequado para o ensaio de formação de esferas em HT-29 e Caco-2 foi o M1 que consistia em: DMEM/F12 Ham suplementado com 2% de B27, 0,5 mg/mL de hidrocortisona, 20 ng/mL de FGF, 20 ng/mL de EGF e 1% de penicilina/estreptomicina. A melhor concentração de células foi 4×10^3 células/poço. Além disso, HT-29 e Caco-2 diferem na sua capacidade de formação de esferas de modo que HT-29 necessita de 26 dias para estabelecer as esferas, enquanto Caco-2 é capaz de fazer isso em apenas 14 dias.

5.2. Identificação dos melhores genes-referência para as análises por RT-qPCR.

Sabendo que a comparação entre uma condição 2D (parental) e 3D (esfera) poderia afetar a expressão de genes-referência, foi realizado um teste com 10 desses genes a fim de identificar aqueles que apresentavam menor variação de Cq (do inglês *Quantification Cycle*) entre as condições parental e esfera. Ao mesmo tempo, eles também deveriam ser amplamente expressos na célula. A tabela 5.1 mostra a média aritmética dos Cqs da condição parental e esfera e a diferença entre esses Cqs obtidos em Caco-2. Conforme é possível observar, existe uma grande variação no Cq da condição parental e esfera para os genes-referência que são mais amplamente usados, como *ACTB* e *18S*, os quais apresentam diferenças de

12,61 e 18,06. Dentre os que apresentaram menor variação de expressão, estão *UBE2D2*, *PGK1* e *PPIA* com diferenças de 2,15; 2,4 e 2,6 respectivamente. Considerando que *UBE2D2* e *PPIA* foram mais expressos em ambas as condições do que *PGK1* eles foram os genes-referência escolhidos para as análises em Caco-2. No entanto, *UBE2D2* não apresentou amplificação em HT-29 e, portanto, para essa linhagem as análises foram realizadas apenas com o gene referência *PPIA*.

Tabela 5.1: Diferença entre as médias aritméticas dos Cqs de genes-referência entre as condições parental e esfera.

Alvo	Média aritmética dos Cqs		Diferença entre Cq Parental e Cq Esfera
	Parental	Esfera	
<i>ACTB</i>	31,59	18,98	12,61
<i>18S</i>	29,40	11,34	18,06
<i>B2M</i>	21,33	24,53	3,2
<i>PGK1</i>	28,62	26,22	2,4
<i>GAPDH</i>	26,06	22,96	3,1
<i>ZUP1</i>	34,37	25,97	8,4
<i>LRRC2</i>	32,55	36,46	3,91
<i>UBE2D2</i>	19,53	21,68	2,15
<i>PPIA</i>	19,04	21,64	2,6
<i>GUSB</i>	33,63	30,64	2,99

5.3. O ensaio de formação de esferas é capaz de induzir mudanças na expressão gênica de marcadores de CTTs em linhagens celulares de câncer colorretal.

Uma vez estabelecidos os genes-referência mais adequados para as análises de expressão gênica, era preciso confirmar se as esferas geradas usando as linhagens celulares Caco-2 e HT-29 realmente consistiam numa população enriquecida para células tronco tumorais. Para isso, cinco diferentes marcadores de CTTs (*ALDH1*, *CD24*, *CD44*, *PROM1* e *LGR5*) tiveram seus níveis de expressão avaliados por RT-qPCR. A figura 5.9 ilustra a variação na expressão desses marcadores em Caco-2 (a) e HT-29 (b). Independente da linhagem avaliada, os consideráveis desvios padrão evidenciam a grande variabilidade gerada pelo ensaio de formação de esferas. Ainda assim, foi possível observar em Caco-2 uma redução estatisticamente significativa na expressão de *CD24*. Já para HT-29 foi possível

observar aumento estatisticamente significativo na expressão de *ALDH1*, *LGR5* e *PROM1*.

Conjuntamente, estes dados sugerem que células Caco-2 submetidas ao ensaio de formação de esferas tem menor expressão do marcador *CD24*. Por outro lado, em HT-29 o ensaio de formação de esferas resulta em células com expressão aumentada de *ALDH1*, *LGR5* e *PROM1*.

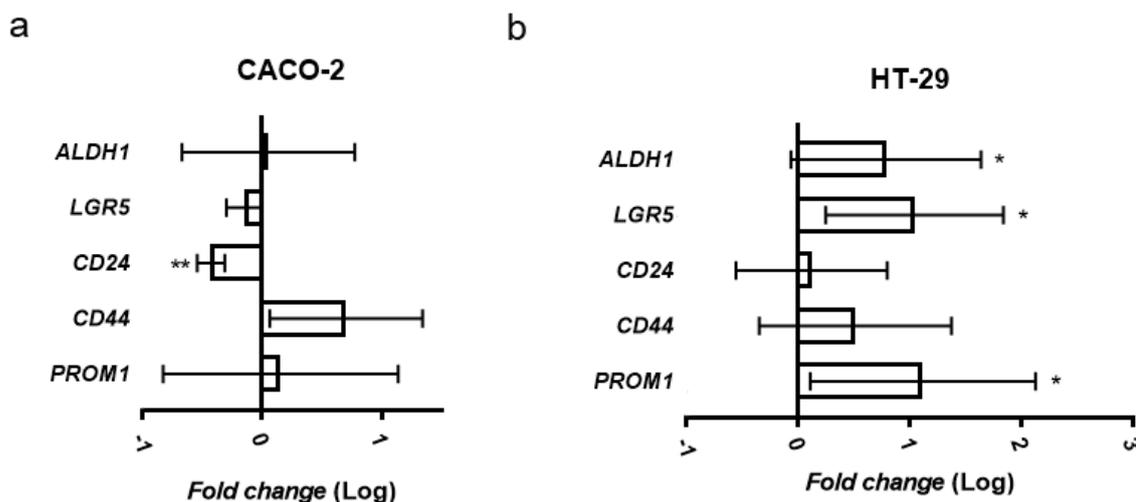


Figura 5.9: Variações na expressão dos marcadores de CTTs em Caco-2 e HT-29 após o ensaio de formação de esferas. As linhagens Caco-2 (a) e HT-29 (b) foram submetidas ao ensaio de formação de esferas e após 14 dias e 26 dias, respectivamente, tiveram seu RNA extraído, c-DNA sintetizado e foram submetidas ao ensaio de RT-qPCR para a análise da variação da expressão gênica dos marcadores de CTTs *ALDH1*, *LGR5*, *CD24*, *CD44* e *PROM1* em comparação com a da condição parental. Gráficos de barras mostram os valores de *fold change* em escala logarítmica da expressão relativa de mRNA obtida pela fórmula $2^{-\Delta\Delta CT}$. Resultados de no mínimo quatro experimentos independentes para Caco-2 e de no mínimo sete experimentos independentes para HT-29. Barras de erro apontam o desvio padrão. * $P < 0.05$ e ** $P < 0.01$.

5.4. Efeito do ensaio de formação de esferas na expressão de glicogenes em linhagens celulares de câncer colorretal.

Tendo estabelecido que o ensaio de formação de esferas gerava alterações na expressão de marcadores de CTTs em Caco-2 e HT-29, o passo seguinte foi investigar se essa condição levava a variações na expressão de genes que codificam enzimas relevantes para a glicosilação. Para isso, foram avaliados 8 glicogenes: *B3GNT2*, *B4GALT1*, *GFPT1*, *GFPT2*, *MGAT5*, *OGA*, *OGT* e *ST3GAL1* (Figura 5.10). Do mesmo modo que ocorreu com os marcadores de CTTs, a variação na expressão dos glicogenes foi grande o que pode ser percebido pelos amplos desvios padrão. Em Caco-2 isso representou uma limitação para o estabelecimento de qualquer conclusão em relação à expressão dos glicogenes. No

entanto, para HT-29, houve aumento estatisticamente significativo na expressão de *B3GNT2*, *OGT* e *OGA*.

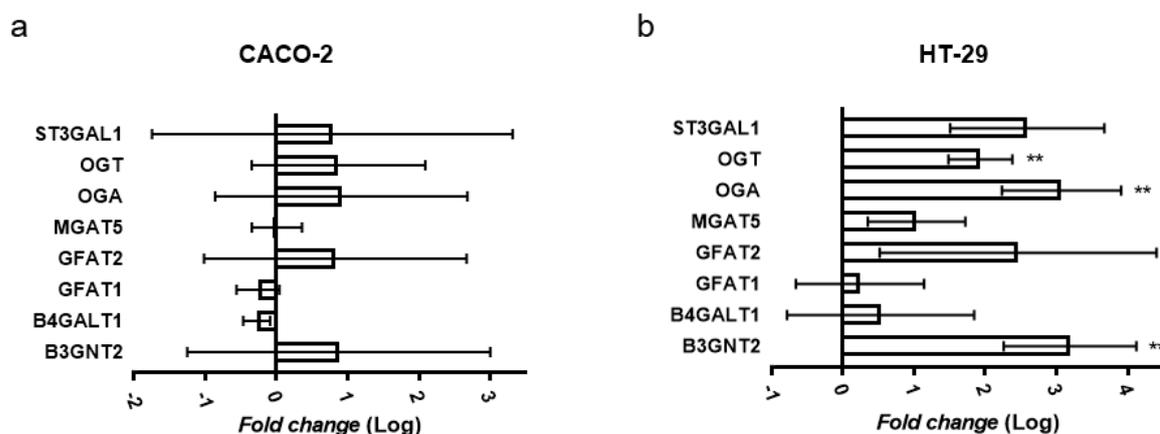


Figura 5.10: Variações na expressão de glicogenes em Caco-2 e HT-29 submetidas ao ensaio de formação de esferas. As linhagens Caco-2 (a) e HT-29 (b) foram submetidas ao ensaio de formação de esferas e após 14 dias e 26 dias, respectivamente, tiveram seu RNA extraído, c-DNA sintetizado e foram submetidas ao ensaio de RT-qPCR para a análise da variação da expressão gênica dos glicogenes *B3GNT2*, *B4GALT1*, *GFPT1*, *GFPT2*, *MGAT5*, *OGA*, *OGT* e *ST3GAL1*. Gráficos de barras mostram os valores de *fold change* em escala logarítmica da expressão relativa de mRNA obtida pela fórmula $2^{-\Delta\Delta CT}$. Resultados de três ou mais experimentos independentes. ** $P < 0.01$.

5.5. Investigação *in silico* da correlação linear entre a expressão gênica dos marcadores de CTTs e a expressão de glicogenes.

Sabendo que em HT-29 houve um aumento estatisticamente significativo na expressão de *B3GNT2*, *OGA* e *OGT* em esferas que apresentavam também níveis aumentados da expressão de três marcadores de CTTs (*ALDH1*, *LGR5* e *PROM1*), foi investigada a possibilidade de haver uma correlação entre a expressão desses glicogenes e a expressão de marcadores de CTTs em amostras de carcinoma de cólon e reto. Para isso, foram realizadas análises *in silico* de correlação linear através da plataforma GEPIA, que utiliza dados depositados no banco COAD (de amostras de carcinoma de cólon) e READ (de amostras de carcinoma de reto) do TCGA (Tabelas 5.2 e 5.3). Conforme ilustra a figura 5.11, para as amostras de carcinoma de cólon existe uma correlação positiva entre *CD44* x *OGA*, *CD44* x *B3GNT2* e *LGR5* x *OGA*. Para as amostras de carcinoma de reto (Figura 5.12), a correlação positiva se manteve na comparação entre *CD44* x *OGA* e *LGR5* x *OGA* indicando que esses genes têm comportamentos similares tanto em carcinoma de cólon quanto de reto, no entanto, exclusivamente nos carcinomas retais, *CD44* x *OGT* e *CD24* x *B3GNT2* correlacionaram-se positivamente.

Esses resultados indicam que podem existir mecanismos relacionados com o estado tronco que envolvam alterações no perfil de glicosilação das CTTs.

Tabela 5.2: Correlações lineares analisadas entre a expressão de marcadores de CTTs e glicogenes em amostras de câncer de cólon indicando o coeficiente de correlação e o valor de *P*.

Correlações testadas	Coefficiente de correlação (r)	P valor
ALDH1 X B3GNT2	-0,048	0,43
ALDH1 X OGA	-0,035	0,56
ALDH1 X OGT	-0,07	0,25
CD24 X B3GNT2	-0,076	0,21
CD24 X OGA	0,093	0,12
CD24 X OGT	0,057	0,34
CD44 X OGT	0,016	0,79
CD133 X B3GNT2	0,0023	0,97
CD133 X OGA	0,099	0,1
CD133 X OGT	-0,075	0,21
LGR5 X B3GNT2	-0,071	0,24
LGR5 X OGT	0,0079	0,9

Tabela 5.3: Correlações lineares analisadas entre a expressão de marcadores de CTTs e glicogenes em amostras de câncer de reto indicando o coeficiente de correlação e o valor de *P*.

Correlações testadas	Coefficiente de correlação (r)	P valor
ALDH1 X B3GNT2	-0,15	0,17
ALDH1 X OGA	0,049	0,64
ALDH1 X OGT	-0,11	0,28
CD24 X OGA	0,034	0,75
CD24 X OGT	0,045	0,67
CD44 X B3GNT2	0,13	0,2
CD133 X B3GNT2	0,03	0,78
CD133 X OGA	-0,013	0,9
CD133 X OGT	-0,017	0,87
LGR5 X B3GNT2	-0,14	0,19
LGR5 X OGT	0,17	0,11

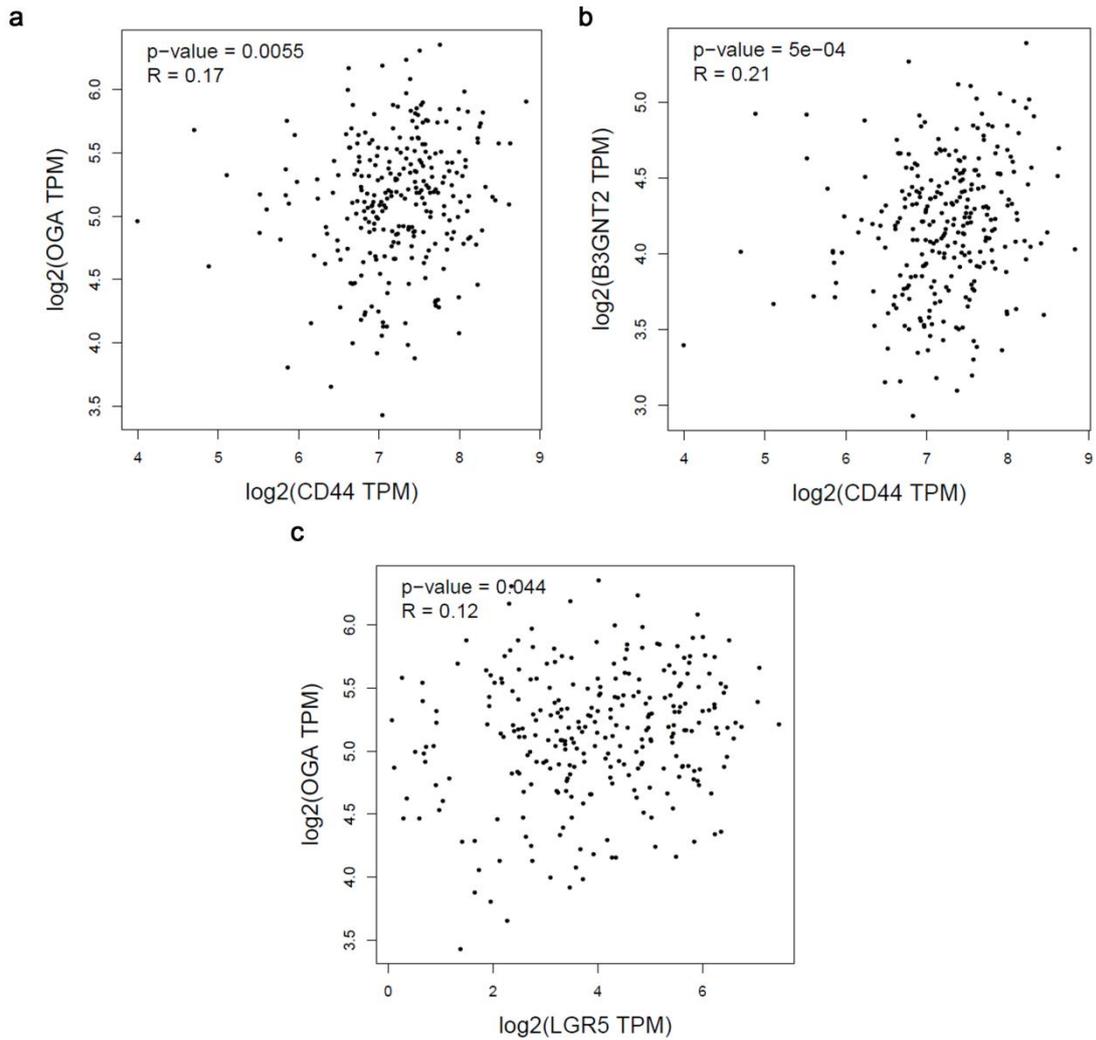


Figura 5.11: Correlação linear entre a expressão de marcadores de CTTs e glicogênes em amostras de carcinoma de cólon. Dados de 275 amostras oriundas do banco COAD do TCGA foram utilizadas para estabelecer correlações entre a expressão de marcadores de CTTs e glicogênes. (a) Indica correlação entre *CD44* e *OGA*, (b) Indica correlação entre *CD44* e *B3GNT2* e (c) Indica correlação entre *LGR5* e *OGA*. R, coeficiente de correlação de Pearson.

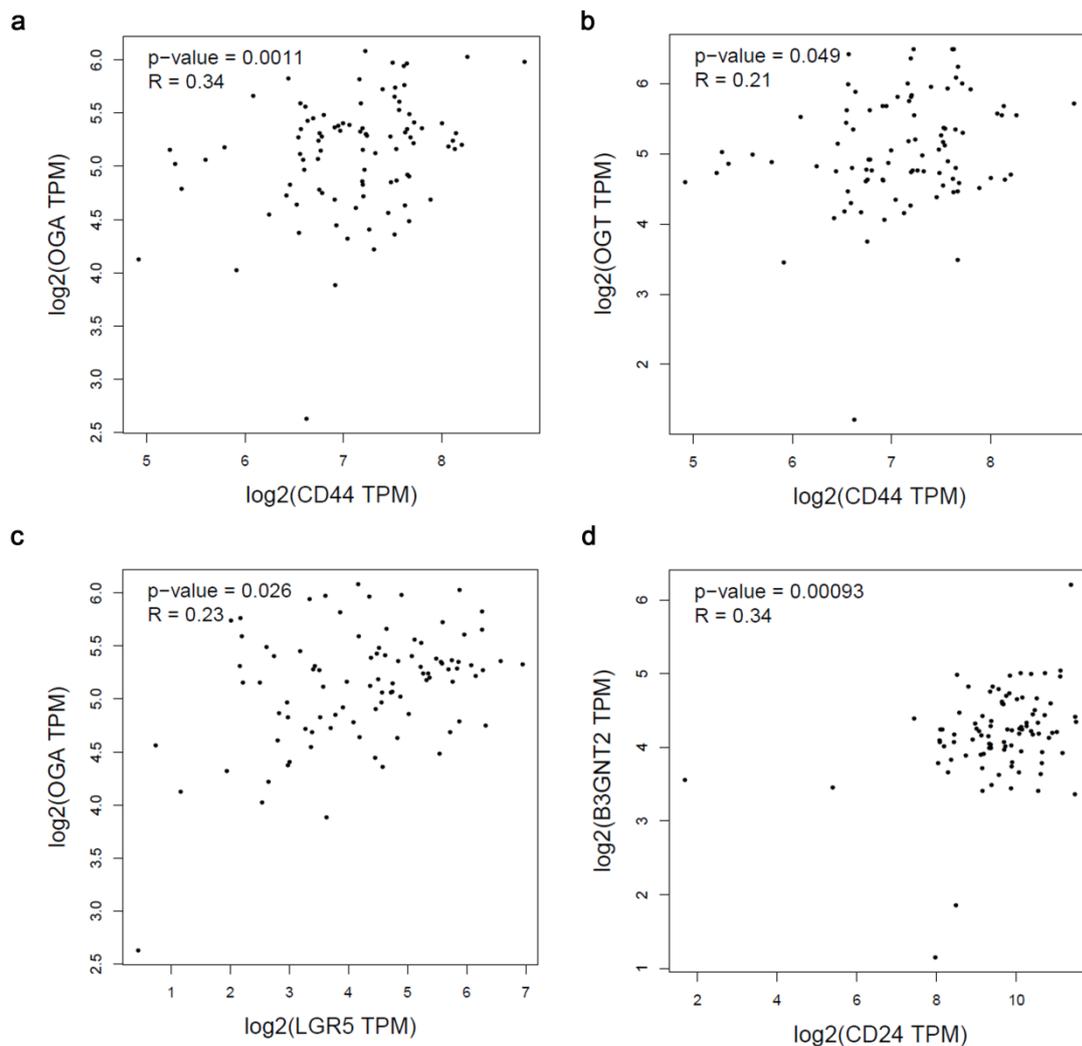


Figura 5.12: Correlação linear entre a expressão de marcadores de CTTs e glicogenes em amostras de carcinoma de reto. Dados de 92 amostras oriundas do banco READ do TCGA foram utilizadas para estabelecer correlações entre a expressão de marcadores de CTTs e glicogenes. (a) Indica correlação entre *CD44* e *OGA*, (b) Indica correlação entre *CD44* e *OGT*, (c) Indica correlação entre *LGR5* e *OGA* e (d) Indica correlação entre *CD24* e *B3GNT2*. R, coeficiente de correlação de Pearson.

5.6. Análises *in silico* do perfil de expressão gênica de marcadores de CTTs.

A fim de avaliar como a expressão dos marcadores de CTTs se comportavam em amostras de pacientes, foram realizadas análises *in silico* através da plataforma UALCAN, que utiliza amostras de tecido adjacente não tumoral e de tumores de cólon e reto depositadas, respectivamente, nos bancos COAD e READ do TCGA. Dessa forma, conforme ilustra a figura 5.13, foi possível verificar que há um aumento significativo da expressão de *LGR5* e *CD44* e uma redução significativa de *CD24* e *ALDH1* em amostras de tumor primário de cólon quando comparadas com amostras adjacentes não tumorais. Quando a mesma comparação foi feita com amostras de reto (Figura 5.14), verificou-se um aumento significativo na expressão de *LGR5*,

CD44 e *PROM1*, e uma redução na expressão de *ALDH1*. Assim, é possível afirmar que tanto em amostras de carcinoma de cólon quanto em amostras de carcinoma de reto há aumento da expressão de *LGR5* e *CD44* e redução da expressão de *ALDH1* quando essas amostras são comparadas com o tecido adjacente não tumoral. No entanto, a redução na expressão de *CD24* é observada apenas nas amostras de carcinoma de cólon, bem como, o aumento na expressão de *PROM1* é observado apenas nas amostras de carcinoma de reto.

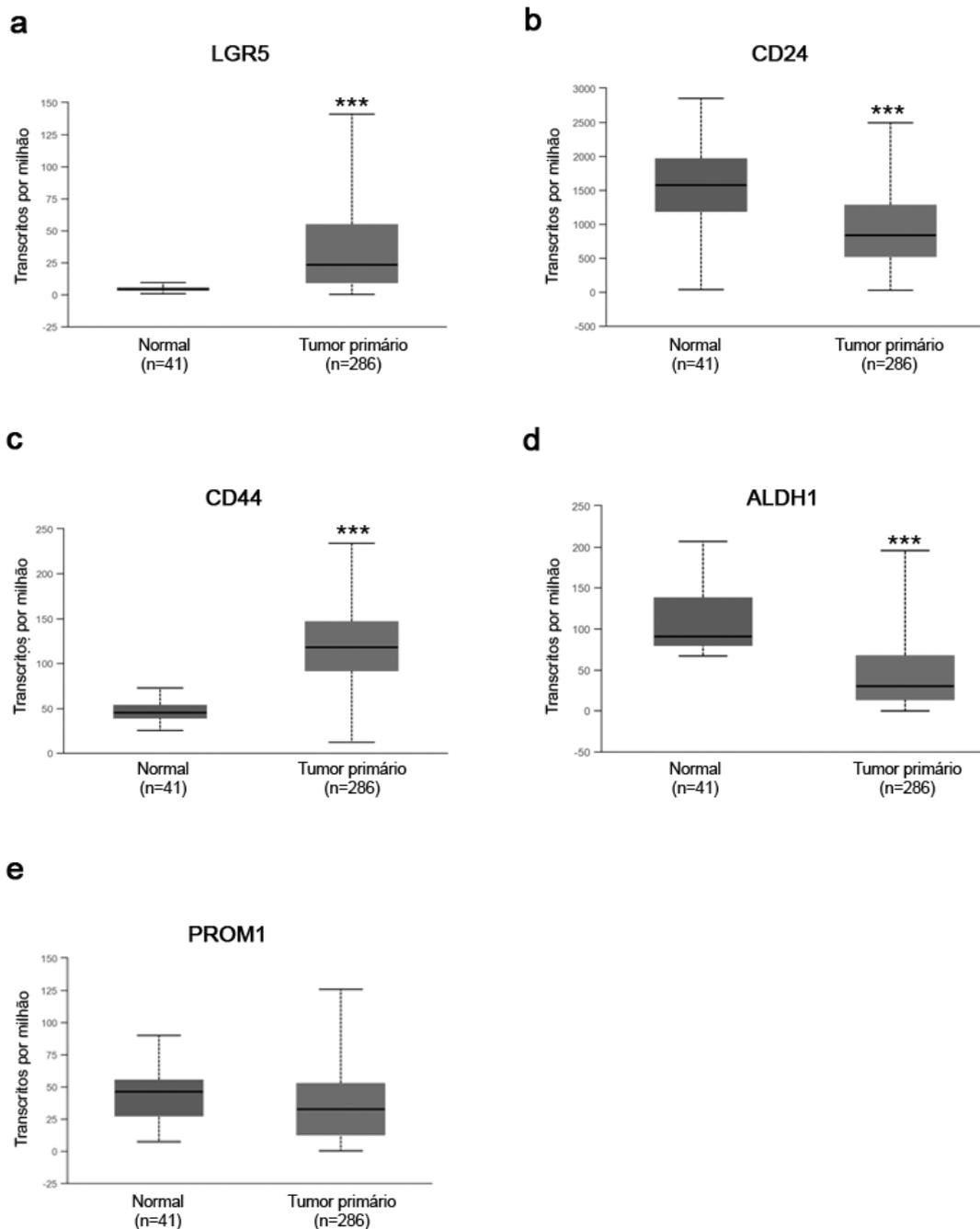


Figura 5.13: Diferenças na expressão dos marcadores de CTTs entre amostras normais adjacentes e de tumor primário de cólon. Amostras adjacentes não tumorais e de tumor primário depositadas no banco de dados COAD do TCGA foram utilizadas para comparar a expressão

diferencial de (a) *LGR5*, (b) *CD24*, (c) *CD44*, (d) *ALDH1* e (e) *PROM1*. Os gráficos representam a mediana e a diferença interquartis e o número de amostras avaliadas em cada grupo é indicado entre parênteses. *** $P < 0.001$.

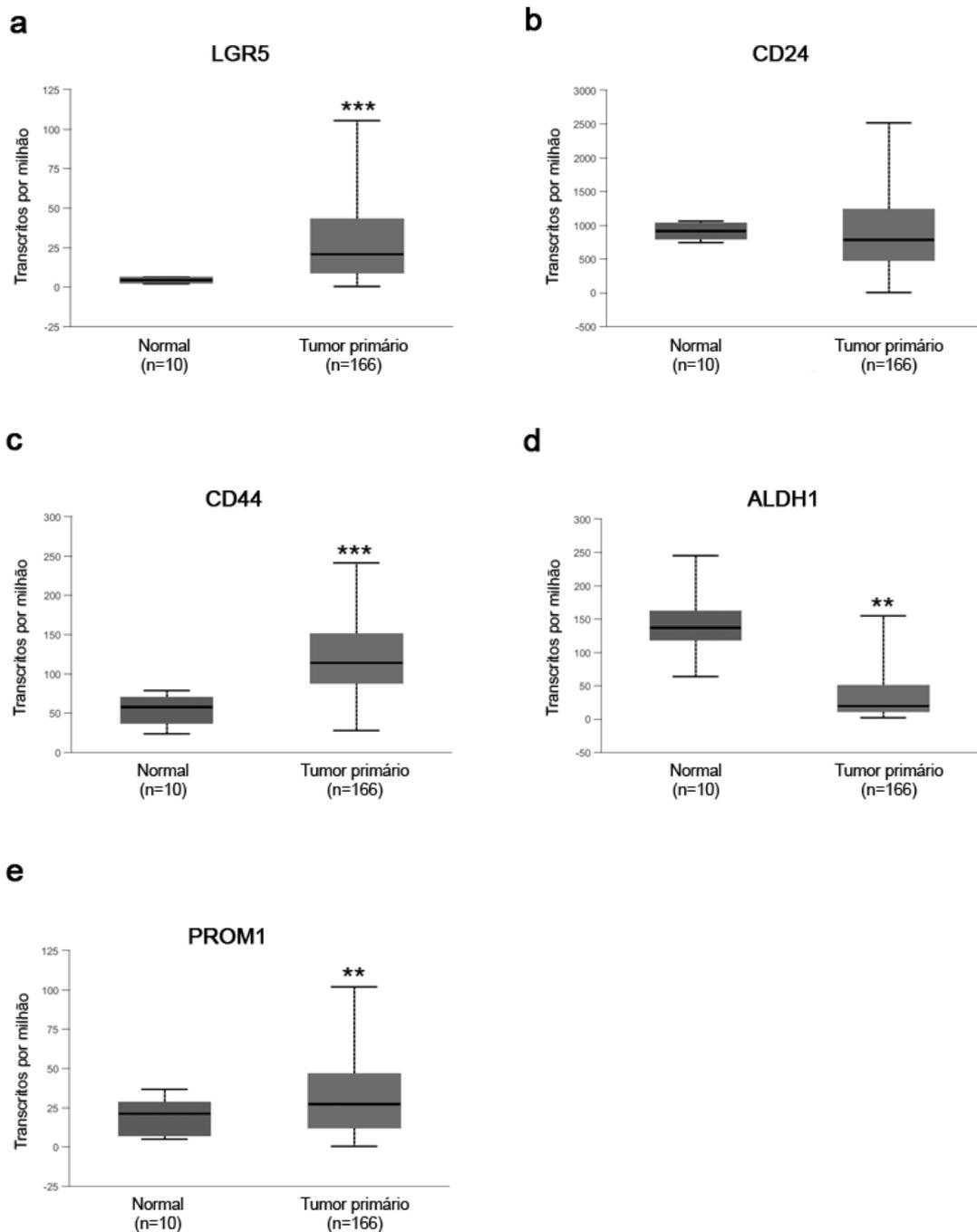


Figura 5.14: Diferenças na expressão dos marcadores de CTTs entre amostras normais adjacentes e de tumor primário de reto. Amostras adjacentes não tumorais e de tumor primário depositadas no banco de dados READ do TCGA foram utilizadas para comparar a expressão diferencial de (a) *LGR5*, (b) *CD24*, (c) *CD44*, (d) *ALDH1* e (e) *PROM1*. Os gráficos representam a mediana e a diferença interquartis e o número de amostras avaliadas em cada grupo é indicado entre parênteses. ** $P < 0.01$ e *** $P < 0.001$.

Sabendo que as CTTs podem ter um papel no desenvolvimento de metástases, foram investigadas também flutuações na expressão dos marcadores correspondentes em diferentes estadios, tanto de câncer de cólon quanto de reto. Nestes casos, foram avaliados apenas os marcadores que apresentaram alterações significativas na sua expressão no tecido tumoral em relação ao tecido não tumoral adjacente. A figura 5.15 mostra as análises referentes ao carcinoma de cólon. Para ele, em relação às amostras adjacentes não tumorais, os níveis de expressão de *LGR5* e *CD44* estão aumentados ao longo de toda a progressão da doença. Enquanto o contrário ocorre para *CD24*. Já para *ALDH1*, a redução na expressão passa a ser estatisticamente significativa a partir do estadio II. Ao comparar diferenças entre os estadios, foi observado que a expressão de *CD44* tem uma redução estatisticamente significativa do estadio I para o estadio IV, o que pode indicar que para o processo de metástase distante seja necessária a redução na expressão de *CD44*.

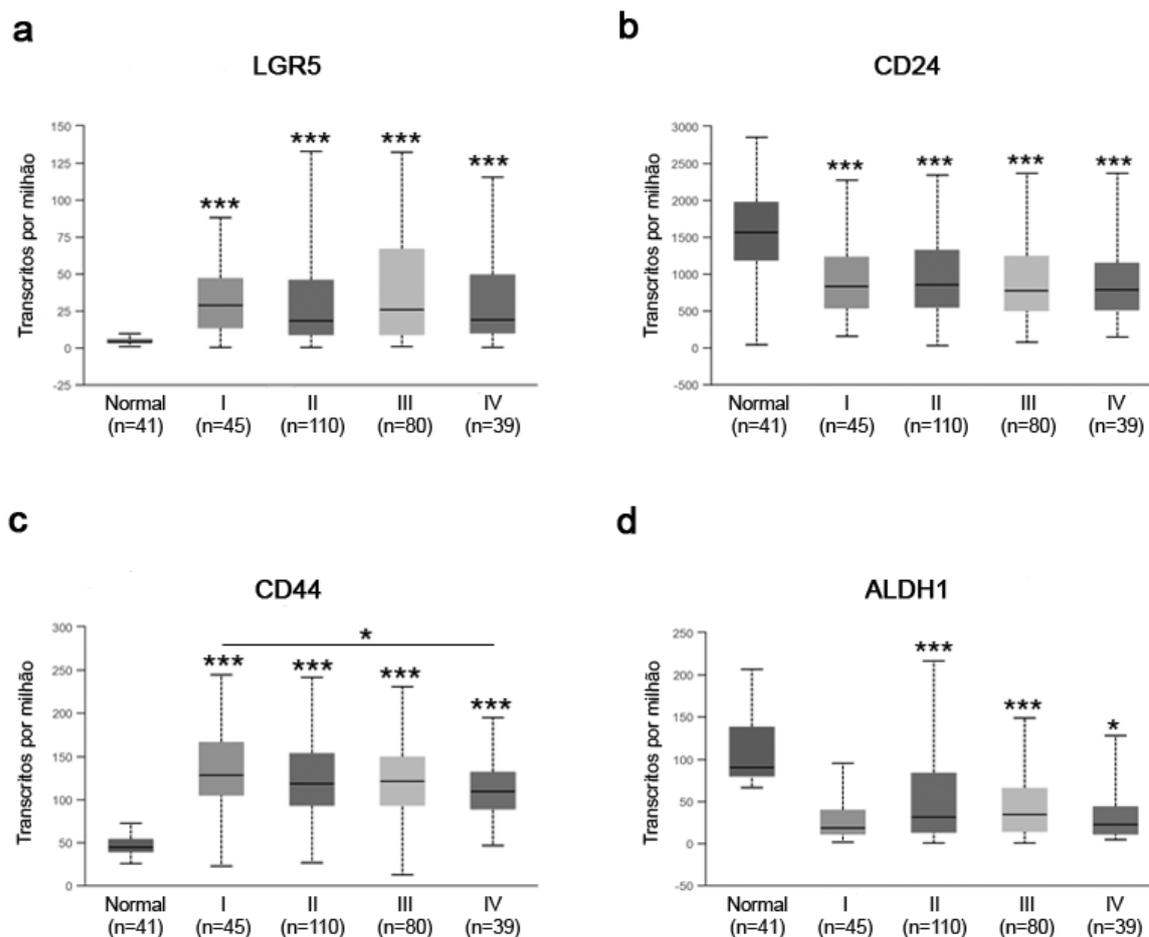


Figura 5.15: Diferenças na expressão de marcadores de CTTs entre os estadios do CCR. Os dados representam a mediana e interquartis de expressão de (a) *LGR5* (b) *CD24* (c) *CD44* (d) *ALDH1* em amostras não tumorais e tumorais separadas em estadios I, II, III e IV ambas depositadas no

banco de dados COAD do TCGA. O número de amostras em cada grupo é indicado entre parênteses. * $P < 0.05$ e *** $P < 0.001$.

Já para carcinoma de reto (Figura 5.16), o mesmo aumento estatisticamente significativo na expressão de *LGR5* e *CD44* é visto nos diferentes estadios da doença, enquanto para *PROM1* ele só ocorre do estadio I ao estadio III, o que pode indicar que *PROM1* é necessário para a progressão da doença, mas não para o processo de metástase distante. Outro fato que também ocorre em amostras de câncer de reto é a redução na expressão de *ALDH1* já a partir do estadio I.

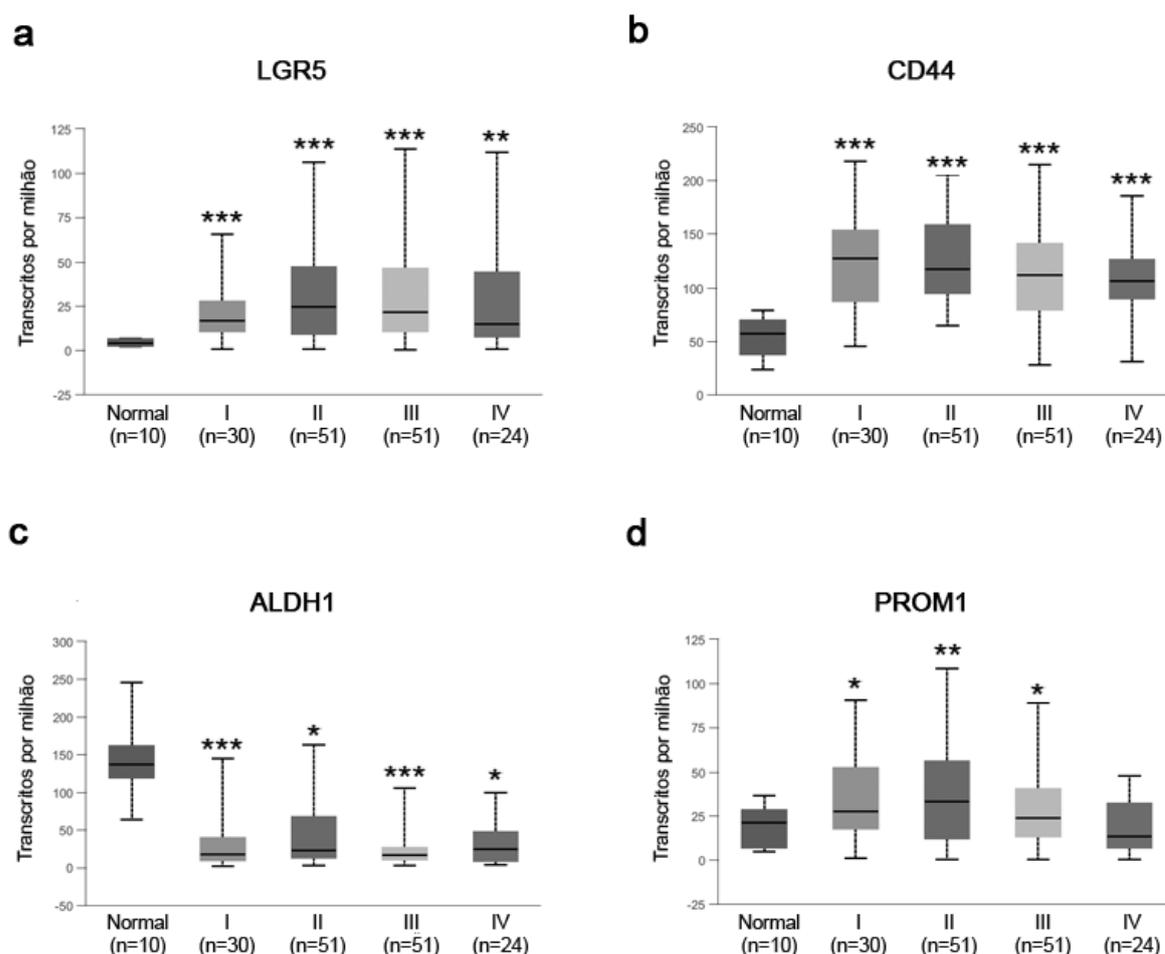


Figura 5.16: Diferenças na expressão de marcadores de CTTs entre os estadios do CCR. Os dados representam a mediana e interquartis de expressão de (a) *LGR5* (b) *CD44* (c) *ALDH1* (d) *PROM1* em amostras não tumorais e tumorais separadas em estadios I, II, III e IV ambas depositadas no banco de dados READ do TCGA. O número de amostras em cada grupo é indicado entre parênteses. * $P < 0.05$, ** $P < 0.01$ e *** $P < 0.001$.

Dessa forma, é possível afirmar que seja no tumor no cólon ou reto há aumento da expressão de *LGR5*, *CD44* e redução de *ALDH1*. Particularmente em cólon, há redução na expressão de *CD24* e em reto aumento de *PROM1* quando se

comparam amostras de tumor primário com amostras de tecido adjacente não tumoral. A maioria dos marcadores apresenta aumento ou redução da expressão ao longo dos diferentes estadios, apenas *CD44* que embora apresente expressão aumentada em todos os estadios nas amostras de carcinoma de cólon, sofre redução da expressão entre o estadio I e IV, sugerindo não ser ele essencial para o desenvolvimento de metástases distantes.

Portanto, esses resultados indicam que em amostras de carcinoma de cólon, há aumento da expressão de *LGR5* e *CD44* e redução de *ALDH1* e *CD24*. Além disso, as análises das amostras separadas por estadiamento sugeriram que uma redução na expressão de *CD44* seja necessária para o processo de metástase distante. Por outro lado, para amostras de carcinoma de reto, houve aumento da expressão de *LGR5*, *CD44* e *PROM1* e redução de *ALDH1*. Quando essas amostras foram separadas por estadiamento, os resultados sugeriram um papel de *PROM1* durante a progressão da doença, mas não durante o processo de metástase distante.

5.7. Análises *in silico* do perfil de expressão gênica de glicogenes

Como a expressão de alguns dos marcadores de CTTs variavam entre amostras de tecido adjacente não tumoral e de tumor primário, análises adicionais foram realizadas a fim de saber como se comportavam os glicogenes estudados nesse mesmo cenário. Em amostras de câncer de cólon, foi observado aumento estatisticamente significativo da expressão de *MGAT5*, *OGA*, *OGT* e *ST3GAL1* e redução da expressão de *B4GALT1*, em comparação com amostras de tecido adjacente não tumoral (Figura 5.17).

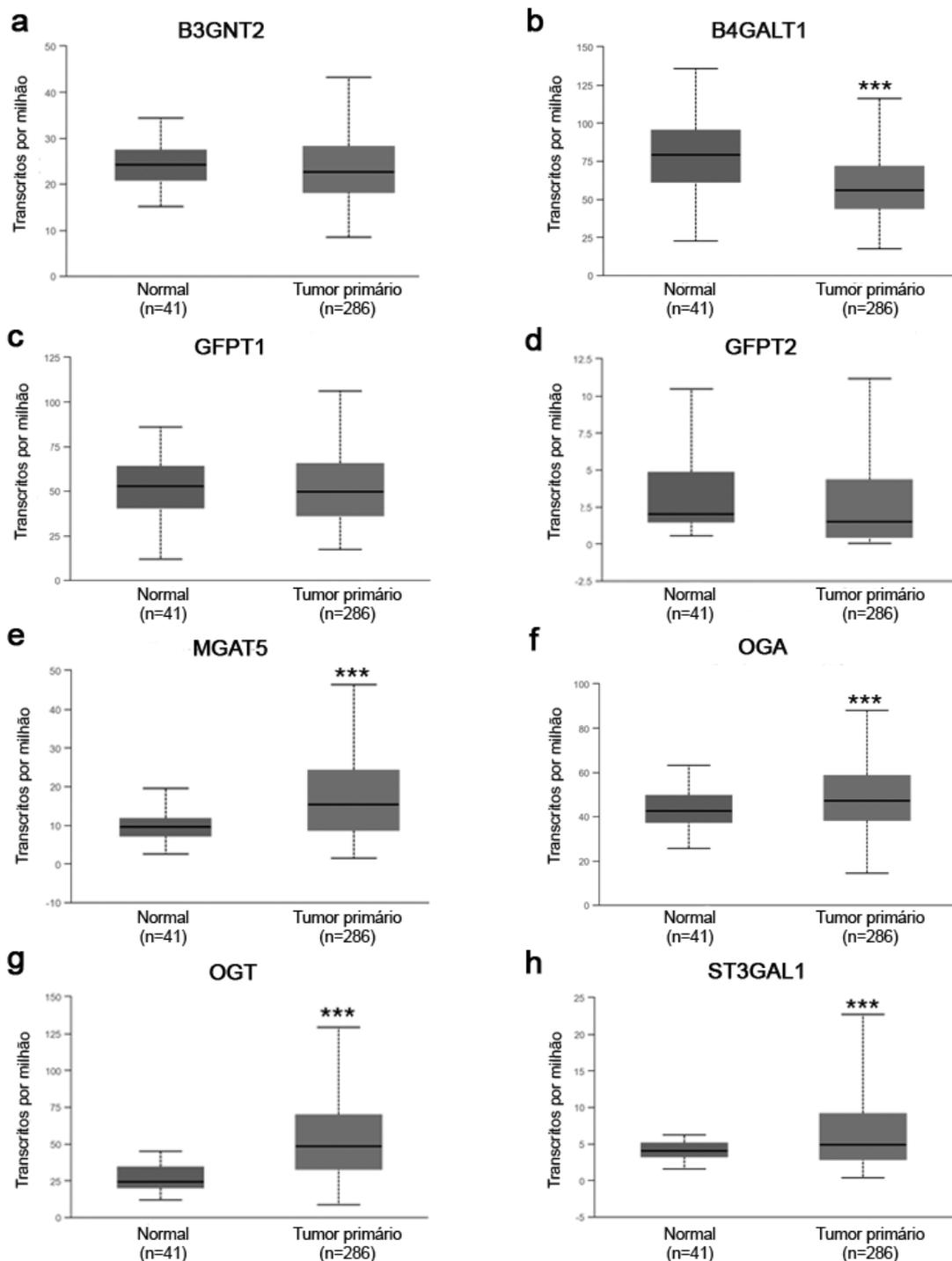


Figura 5.17: Diferenças na expressão de glicogenes entre amostras normais adjacentes e de tumor primário de cólon. Amostras adjacentes não tumorais e de tumor primário depositadas no banco de dados COAD do TCGA foram utilizadas para comparar as diferenças na expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1*. Os gráficos representam a mediana e a diferença interquartis e o número de amostras avaliadas em cada grupo é indicado entre parênteses. *** $P < 0.001$.

Por outro lado, a única alteração que se manteve para os tumores primários de reto foi o aumento na expressão de *OGT*. Além dela, houve redução na

expressão de *B3GNT2* e *ST3GAL1* entre amostras de tumor primário de reto e de tecido não tumoral adjacente (Figura 5.18).

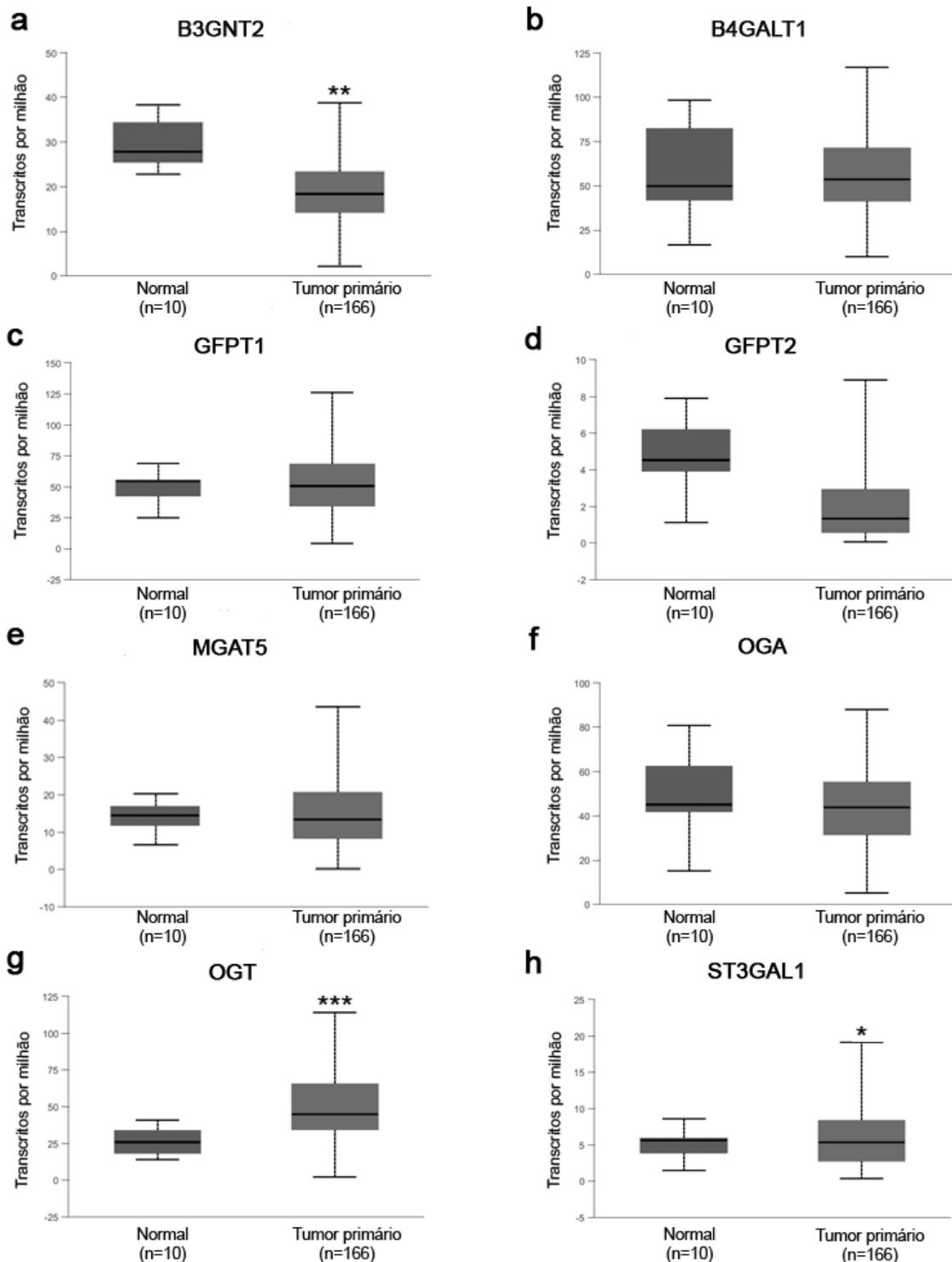


Figura 5.18: Diferenças na expressão de glicogenes entre amostras normais adjacentes e de tumor primário de reto. Amostras adjacentes não tumorais e de tumor primário depositadas no banco de dados READ do TCGA foram utilizadas para comparar as diferenças na expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1*. Os gráficos representam a mediana e a diferença interquartil e o número de amostras avaliadas em cada grupo é indicado entre parênteses. * $P < 0.05$, ** $P < 0.01$ e *** $P < 0.001$.

Para saber se as mudanças na expressão desses glicogenes ocorriam em algum momento específico da progressão tumoral, as amostras tumorais foram então separadas por estadiamento e apenas os glicogenes que apresentavam alteração de expressão entre o tecido tumoral e o tecido adjacente não tumoral foram avaliados. Em relação às análises referentes ao carcinoma de cólon, foi verificado que os níveis de expressão de *MGAT5* e *OGT* encontram-se aumentados ao longo de toda a progressão da doença. Por outro lado, a expressão de *B4GALT1* está diminuída em todos os estadios, embora entre o estadio II e III pareça ocorrer uma tendência de aumento. Já a expressão de *OGA* tem aumento significativo em relação ao tecido adjacente não tumoral apenas nos estadios tardios III e IV. Verificou-se também que a expressão de *ST3GAL1* aumenta de forma significativa em relação ao tecido não tumoral adjacente do estadio II em diante. Por outro lado, fazendo uma comparação entre os estadios, observou-se que a expressão de *OGA* tem um aumento significativo do estadio II para o III, enquanto para *OGT* o mesmo é visto do estadio II para o IV, o que sugere um papel de *OGT* no estabelecimento de metástases distantes. Para *ST3GAL1*, também há aumento estatisticamente significativo quando se compara a expressão no estadio I com os demais (Figura 5.19).

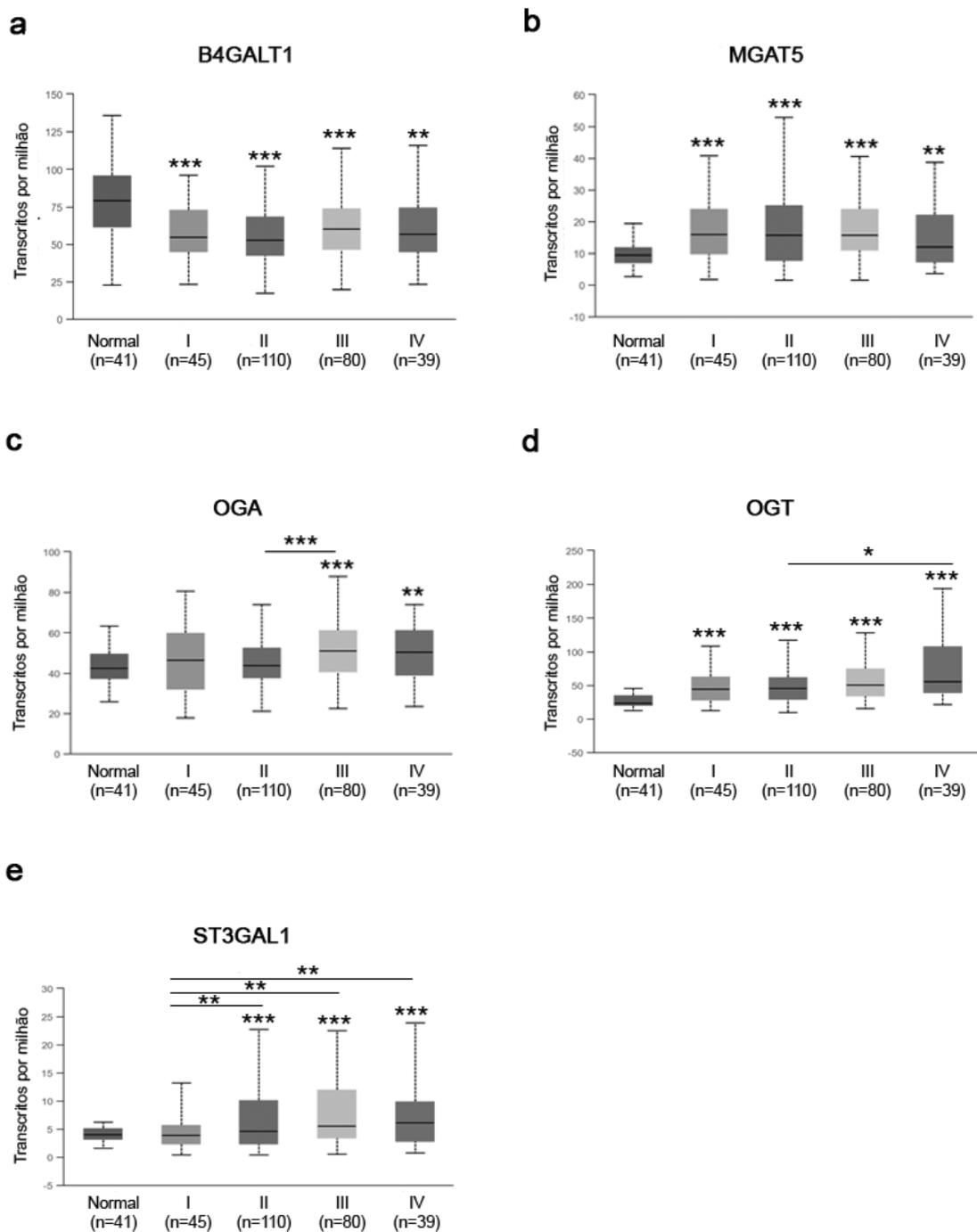


Figura 5.19: Diferenças na expressão de glicogènes entre os estadios do CCR. Os dados representam a mediana e interquartis de expressão dos glicogènes (a) *B4GALT1* (b) *MGAT5* (c) *OGA* (d) *OGT* (e) *ST3GAL1* em amostras não tumorais e tumorais separadas em estadios I, II, III e IV ambas depositadas no banco de dados COAD do TCGA. O número de amostras em cada grupo é indicado entre parênteses. * $P < 0.05$, ** $P < 0.01$ e *** $P < 0.001$.

As análises realizadas com as amostras de reto (Figura 5.20) apontaram redução estatisticamente significativa na expressão de *B3GNT2* do estadio I ao III, aumento na expressão de *OGT* em todos os estadios e pequena redução da expressão de *ST3GAL1* no estadio III, com posterior aumento discreto no estadio IV

o que sugere um possível papel de *ST3GAL1* no desenvolvimento de metástases distantes.

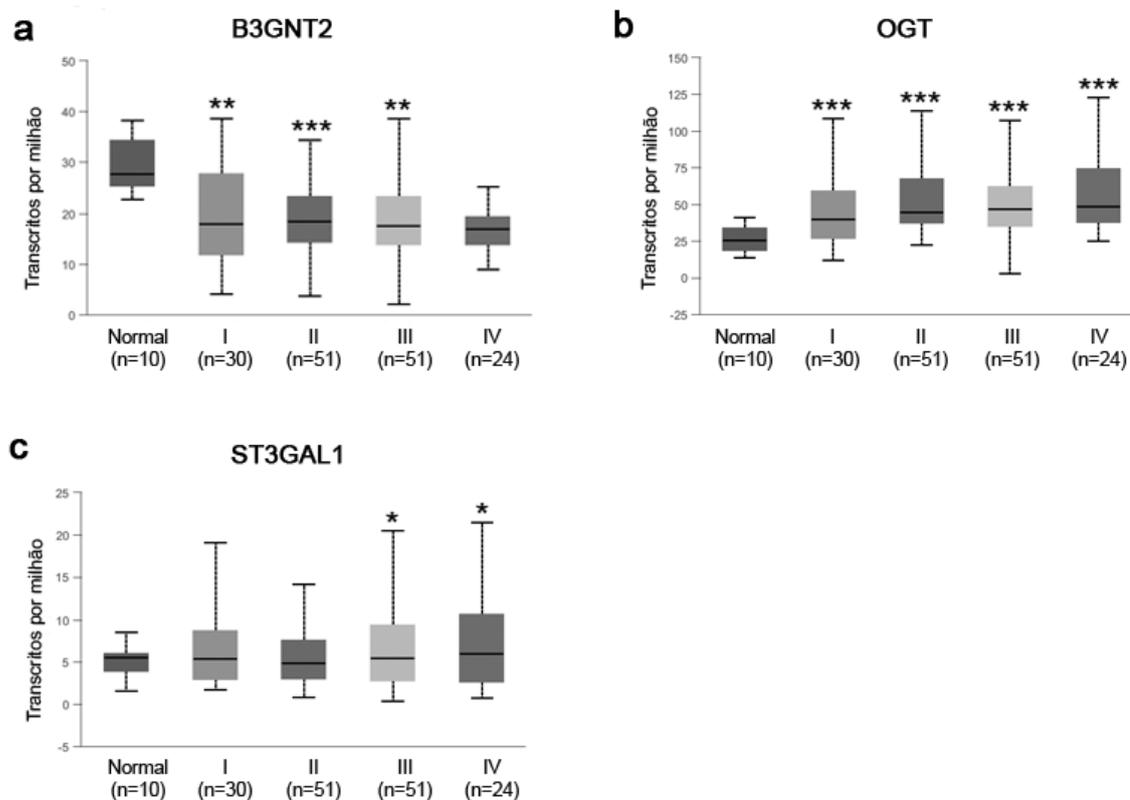


Figura 5.20: Diferenças na expressão de glicogenes entre os estadios do CCR. Os dados representam a mediana e interquartil de expressão dos glicogenes (a) *B3GNT2* (b) *OGT* (c) *ST3GAL1* em amostras não tumorais e tumorais separadas em estadios I, II, III e IV ambas depositadas no banco de dados READ do TCGA. O número de amostras em cada grupo é indicado entre parênteses. * $P < 0.05$, ** $P < 0.01$ e *** $P < 0.001$.

Levando em consideração tanto a tendência de aumento na expressão de *B4GALT1* quanto o aumento estatisticamente significativo na expressão de *OGA* nos estadios II e III para as análises com amostras de cólon, levantou-se a hipótese de que esses glicogenes pudessem estar envolvidos no processo de metástase dos linfonodos, uma vez que no estadio II ainda não há comprometimento linfonodal, o que já é visto no estadio III. Portanto, as amostras foram então classificadas de N0 a N2, de acordo com a quantidade de linfonodos acometidos, onde N0 representa aquelas que não apresentam metástase linfonodal, N1 aquelas que apresentam de 1 a 3 linfonodos regionais acometidos e N2 aquelas que apresentam 4 ou mais linfonodos regionais acometidos (Figura 5.21). Dessa forma, foi possível perceber que há um aumento na expressão de *B4GALT1* entre amostras N0 e N1 indicando

que essa enzima pode ser necessária para o processo de implantação da metástase linfonodal. Por outro lado, a expressão de *OGA* aumenta de N0 para N2 sugerindo um papel de *OGA* na disseminação das metástases linfonodais.

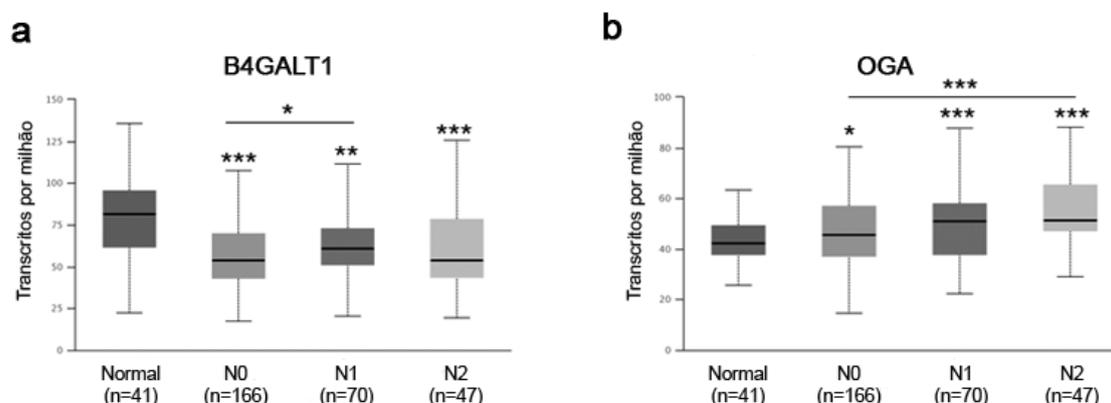


Figura 5.21: Expressão de *B4GALT1* e *OGA* de acordo com o acometimento linfonodal. Amostras depositadas no banco de dados COAD do TCGA foram separadas de acordo com o acometimento linfonodal em N0, N1 ou N2 onde N0 representa amostras que não apresentam metástase linfonodal, N1 aquelas que apresentam de 1 a 3 linfonodos regionais acometidos e N2 as que apresentam 4 ou mais linfonodos regionais acometidos. Feito isso, os níveis de expressão de (a) *B4GALT1* e (b) *OGA* foram analisados. Os gráficos representam a mediana e interquartis. O número de amostras em cada grupo é indicado entre parênteses. * $P < 0.05$, ** $P < 0.01$ e *** $P < 0.001$

Ao contrário do cenário com grandes sobreposições que se observa para a expressão dos marcadores de CTTs entre tumores de cólon e reto, quando se trata da expressão de glicogenes, apenas *OGT* teve aumento em ambos. Foram particulares dos tumores de cólon o aumento dos níveis de expressão de *MGAT5*, *OGA*, *ST3GAL1* e de *B4GALT1*, por outro lado, particularmente em reto, foi verificada redução na expressão de *B3GNT2* e *ST3GAL1*. Assim como ocorreu para a expressão dos marcadores de CTTs, a maioria das alterações na expressão dos glicogenes se manteve nos diferentes estadios. Vale ressaltar que para as análises de carcinoma de cólon houve aumento na expressão: de *OGT* do estadio II para o IV; de *OGA* do II para o III; além da leve tendência de aumento em *B4GALT1* entre esses mesmos estadios. A variação da expressão observada para o gene *OGT* poderia indicar uma contribuição deste no desenvolvimento da metástase distante. Já as variações de expressão identificadas tanto em *OGA* quanto em *B4GALT1* sugerem um possível papel destes no desenvolvimento de metástase linfonodal. A análise do acometimento linfonodal mostrando aumento dos níveis de expressão de *B4GALT1* de N0 para N1 e de *OGA* de N0 para N2 corrobora para a hipótese de que esses glicogenes participem de alguma forma do desenvolvimento de metástases linfonodais e de sua disseminação, respectivamente.

Resumidamente, os resultados indicam que, para amostras de carcinoma de cólon, ocorreu aumento da expressão de *MGAT5*, *OGA*, *OGT* e *ST3GAL1* e redução de *B4GALT1*. As análises por estadiamento sugerem um papel de *OGT* no estabelecimento de metástases distantes, enquanto as análises de acometimento linfonodal sugerem que *B4GALT1* possa ter algum envolvimento neste processo. Os achados sugerem também que *OGA* possa ter algum envolvimento na disseminação das metástases linfonodais. Por outro lado, o que se observou nas amostras de carcinoma de reto foi a redução da expressão de *B3GNT2* e *ST3GAL1* além do aumento de *OGT*. Durante as análises por estadiamento, o perfil observado para *ST3GAL1* levanta a possibilidade de que este gene possa ter alguma influência no desenvolvimento de metástases distantes.

5.8. Avaliação do impacto da expressão de marcadores de CTTs sobre a sobrevida de pacientes com câncer colorretal

Considerando a hipótese de que os pacientes que apresentam maiores níveis de expressão de marcadores de CTTs poderiam ter maior representatividade dessas células na massa tumoral e que isso poderia repercutir no desfecho clínico, foram realizadas, através da plataforma GEPIA, análises de sobrevida global e sobrevida livre de doença por meio da geração de curvas de Kaplan-Meier, utilizando amostras depositadas nos bancos de dados COAD e READ do TCGA. As amostras foram separadas em tercís de expressão, e aquelas pertencentes ao tercil superior (*cut-off* em 66%) foram consideradas com alta expressão (*high*), enquanto as pertencentes ao tercil inferior (*cut-off* em 33%) foram consideradas com baixa expressão (*low*). Surpreendentemente, nenhum dos marcadores de CTTs analisados pode prever a sobrevida global (Figuras 5.22 e 5.24) ou a sobrevida livre de doença (Figuras 5.23 e 5.25) dos pacientes, independentemente da localização do tumor (cólon ou reto). Dessa forma, esses resultados demonstram a incapacidade dos marcadores de CTTs em apresentar qualquer relação com o prognóstico dos pacientes com CCR. Essa incapacidade poderia estar relacionada com o fato das análises terem sido realizadas tomando como base os dados de expressão gênica e não os dados dos níveis proteicos. Outra possibilidade poderia estar relacionada ao longo período de seguimento utilizado nas análises.

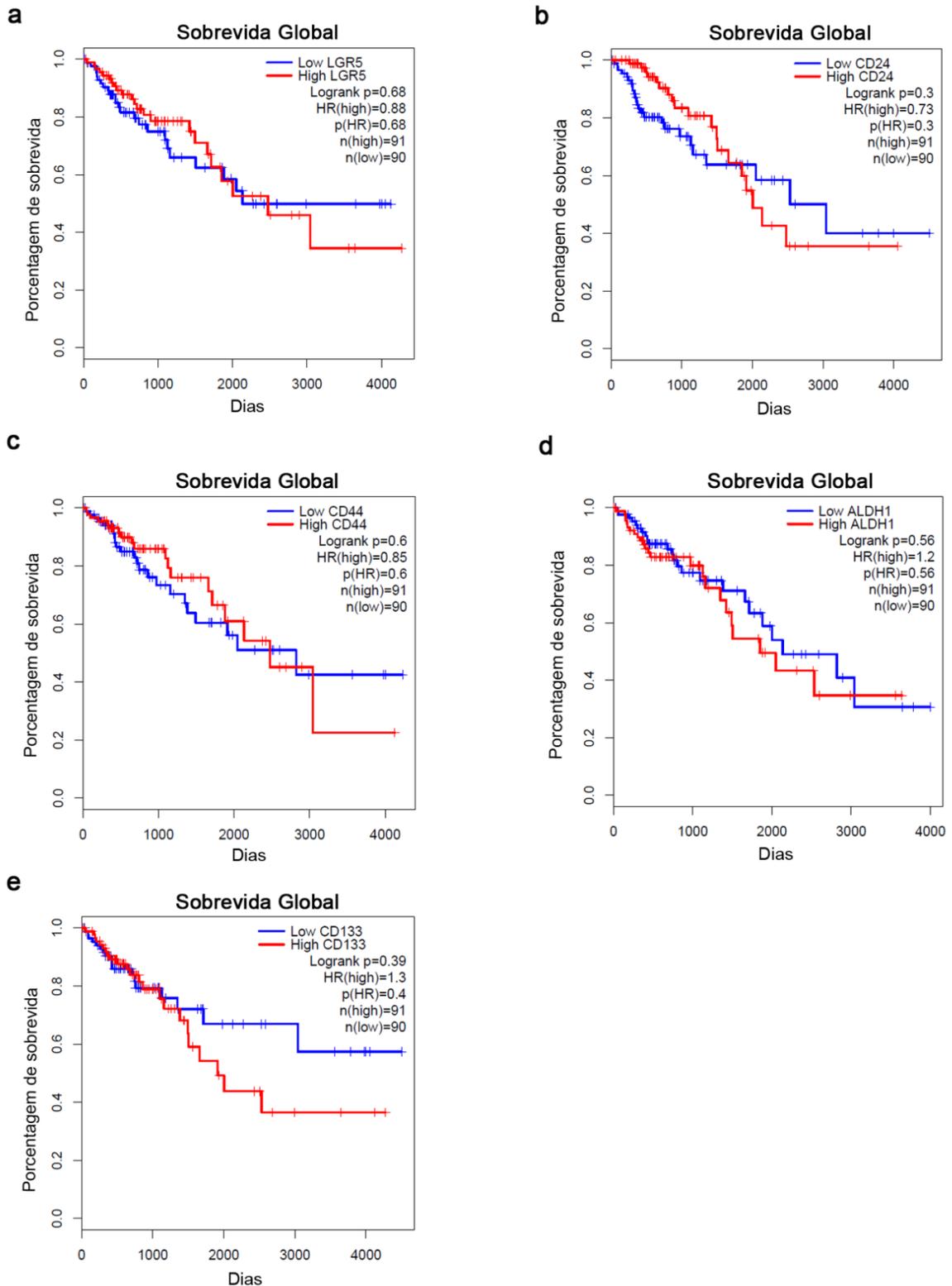


Figura 5.22: Sobrevida global de acordo com os níveis de expressão de diferentes marcadores de CTTs em amostras de câncer de cólon. Amostras depositadas no banco de dados COAD do TCGA foram separadas em tercís de expressão e aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*), enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *LGR5*, (b) *CD24*, (c) *CD44*, (d) *ALDH1* e (e) *CD133* na sobrevivência global dos pacientes. HR, *Hazard Ratio*; $p(\text{HR})$, valor de p da regressão de Cox usada para cálculo do HR.

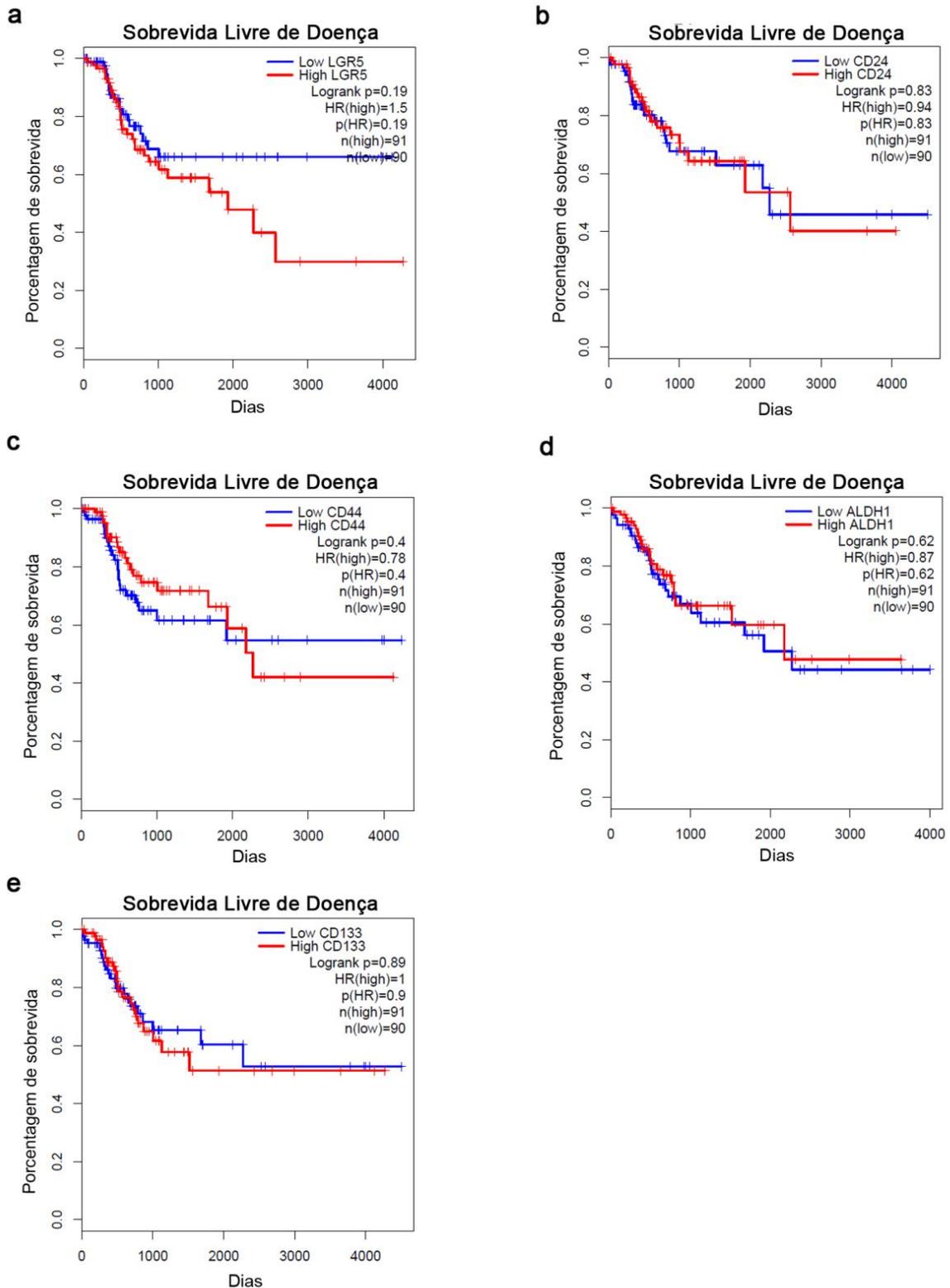


Figura 5.23: Sobrevida livre de doença de acordo com os níveis de expressão de diferentes marcadores de CTTs em amostras de câncer de cólon. Amostras depositadas no banco de dados COAD do TCGA foram separadas em tercís de expressão e aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *LGR5*, (b) *CD24*, (c) *CD44*, (d) *ALDH1* e (e) *CD133* na sobrevivida livre de doença dos pacientes. HR, *Hazard Ratio*; $p(HR)$, valor de p da regressão de Cox usada para cálculo do HR.

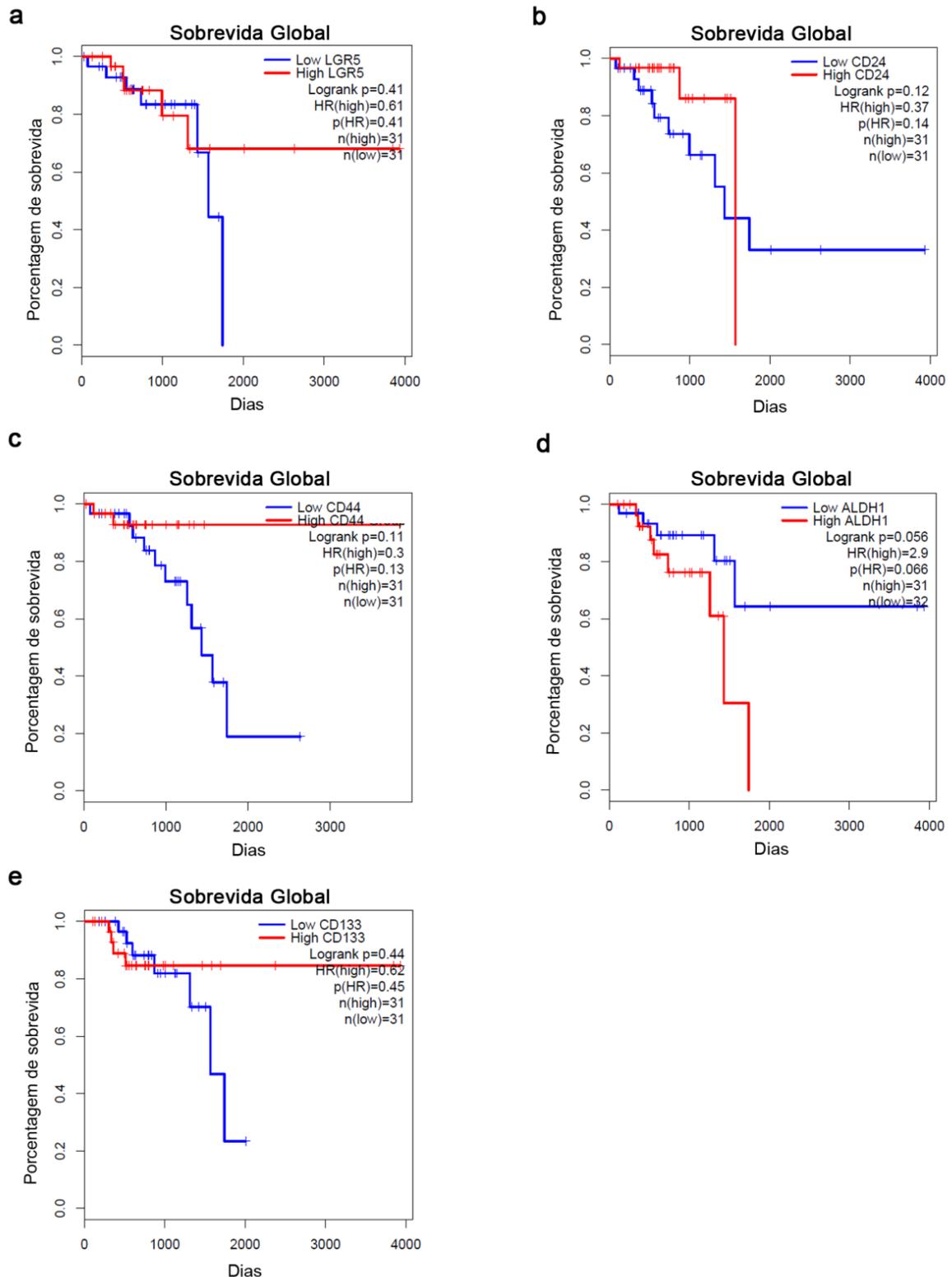


Figura 5.24: Sobrevida global de acordo com os níveis de expressão de diferentes marcadores de CTTs em amostras de câncer de reto. Amostras depositadas no banco de dados READ do TCGA foram separadas em tercís de expressão e aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *LGR5*, (b) *CD24*, (c) *CD44*, (d) *ALDH1* e (e) *CD133* na sobrevida global dos pacientes. Nenhum dos marcadores foi capaz de prever a sobrevida global em pacientes com câncer de reto. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.

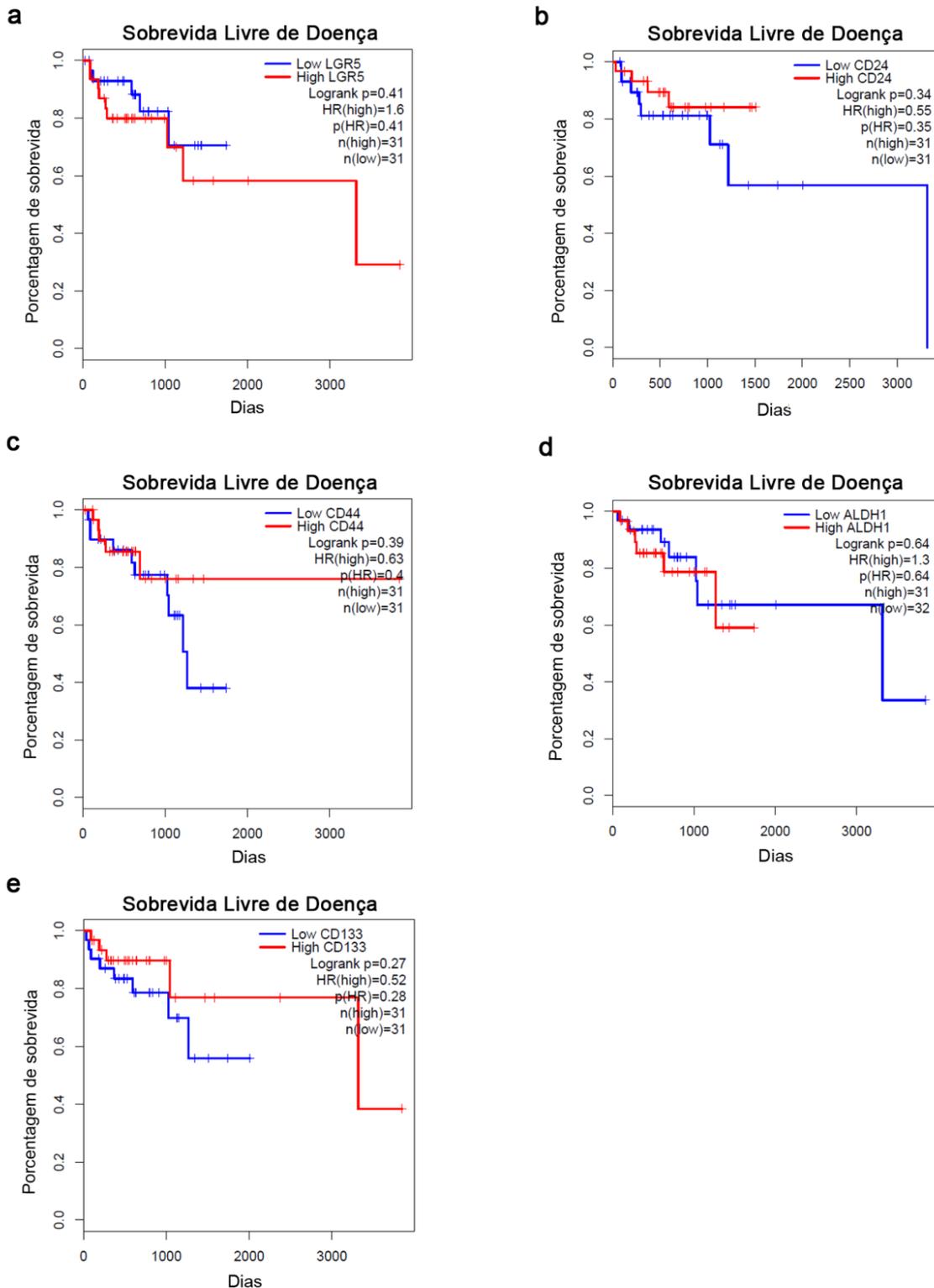
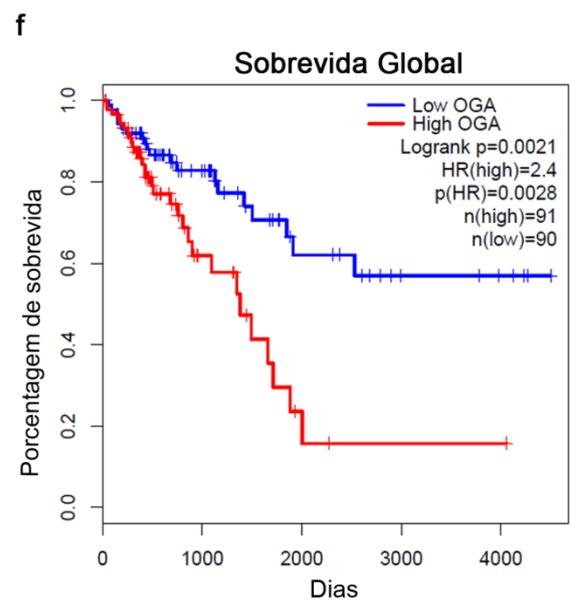
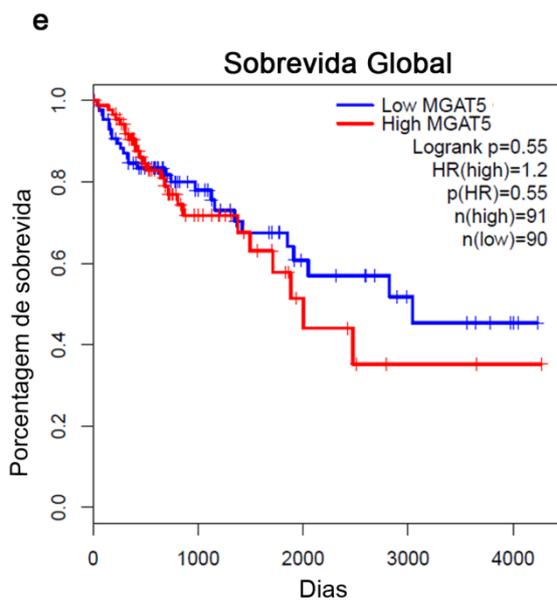
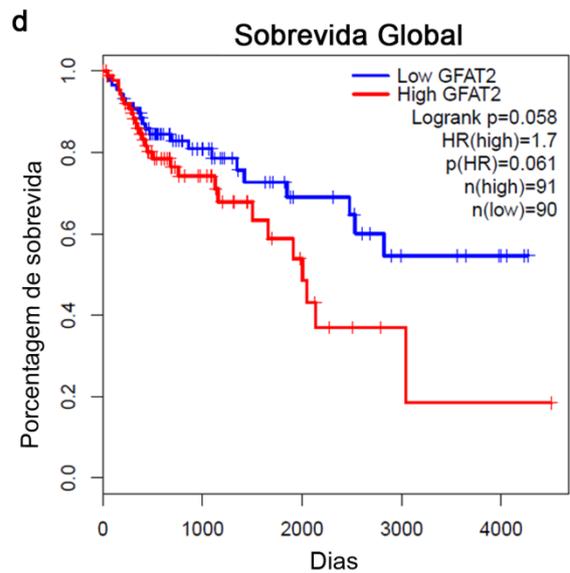
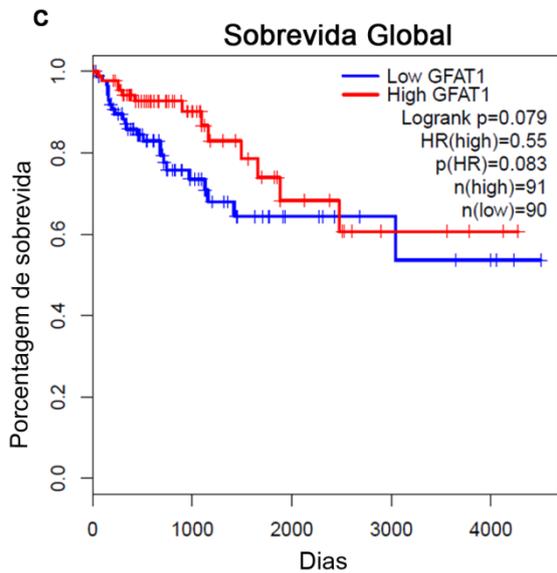
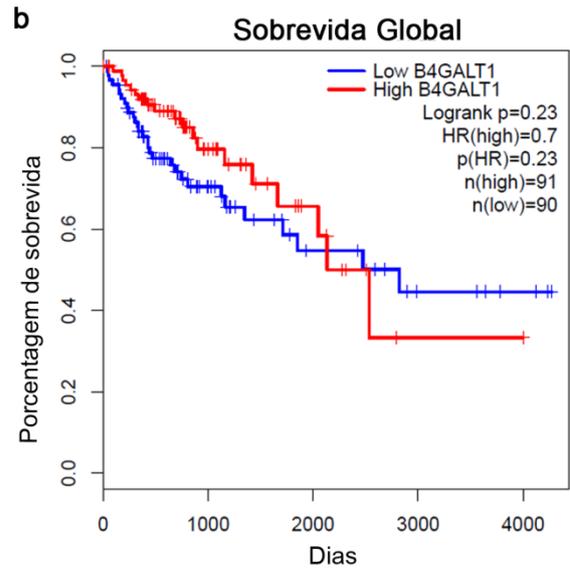
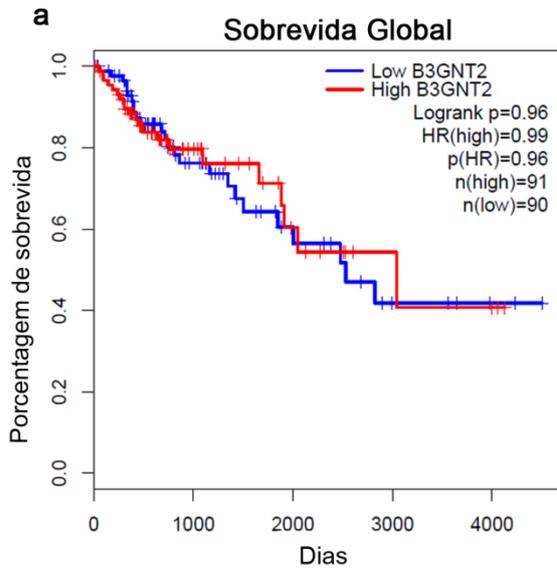


Figura 5.25: Sobrevida livre de doença de acordo com os níveis de expressão de diferentes marcadores de CTTs em amostras de câncer de reto. Amostras depositadas no banco de dados READ do TCGA foram separadas em tercís de expressão e aquelas pertencentes ao tercís superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercís inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *LGR5*, (b) *CD24*, (c) *CD44*, (d) *ALDH1* e (e) *CD133* na sobrevivida livre de doença dos pacientes. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.

5.9. Avaliação do impacto da expressão de glicogenes sobre a sobrevida de pacientes com câncer colorretal

Embora os marcadores de CTTs não tenham, sido capazes de predizer a resposta clínica dos pacientes, neste trabalho foi demonstrado que eles apresentaram correlação positiva com a expressão de alguns glicogenes, assim, admite-se especular que os glicogenes possam de alguma forma regular aspectos fundamentais de subpopulações de CTTs, o que, conseqüentemente, poderia afetar o desfecho da doença. Considerando esta possibilidade, amostras de carcinoma de cólon e reto depositadas no TCGA foram avaliadas de acordo com a expressão dos seguintes glicogenes: *B3GNT2*, *B4GALT1*, *GFAT1*, *GFAT2*, *MGAT5*, *OGA*, *OGT*, *ST3GAL1*. As amostras também foram separadas em tercís de expressão, como realizado para os marcadores de CTTs. A partir daí foi possível estabelecer curvas de sobrevida global e de sobrevida livre de doença, as quais são mostradas nas figuras 5.26 e 5.27, respectivamente. Para as amostras de carcinoma de cólon, a expressão aumentada de *OGA* resultou numa menor sobrevida global e representou um risco de morte 2,4 vezes maior para esse grupo. Por outro lado, pacientes com alta expressão de *MGAT5* apresentaram menor sobrevida livre de doença e um risco duas vezes maior de recorrência que os pacientes com baixa expressão de *MGAT5*. Se os pacientes analisados são estratificados de acordo com seu perfil de regiões microsatélite (MSI ou MSS) os resultados anteriores se mantêm verdadeiros apenas para aqueles que são MSS, porém, neste caso os riscos aumentam, passando a ser 3,5 vezes maior para a morte de pacientes com alta expressão de *OGA* e 3,9 vezes maior para a recorrência em pacientes com alta expressão de *MGAT5* (Figura 5.28). No entanto, entre aqueles que são MSI nenhum dos glicogenes avaliados foi capaz de predizer nem a sobrevida global nem a sobrevida livre de doença. Essa incapacidade persistiu para as amostras de reto, fossem elas estratificadas ou não (Dados apresentados no anexo 1).

Esses resultados demonstram que maiores níveis de expressão de dois dos glicogenes analisados (*MGAT5* e *OGA*) tem valor prognóstico no desfecho clínico de pacientes com carcinoma de cólon.



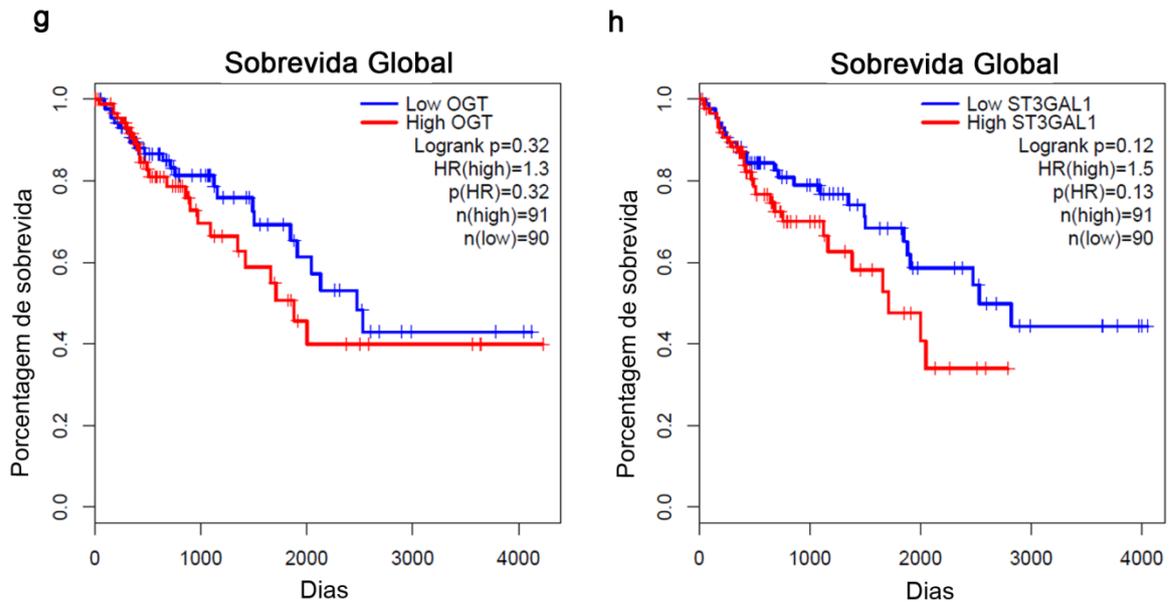
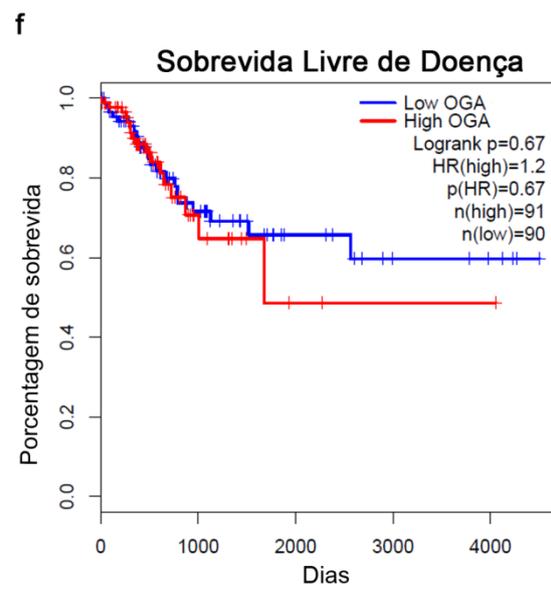
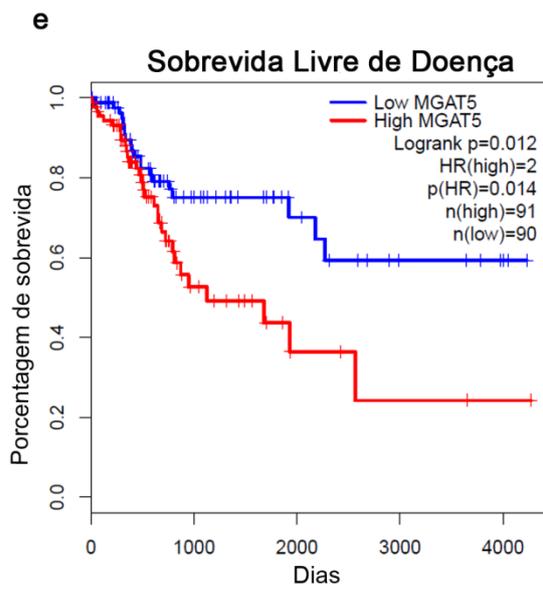
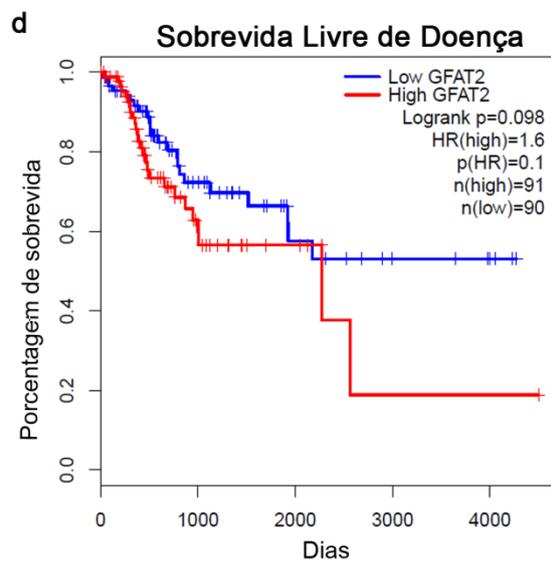
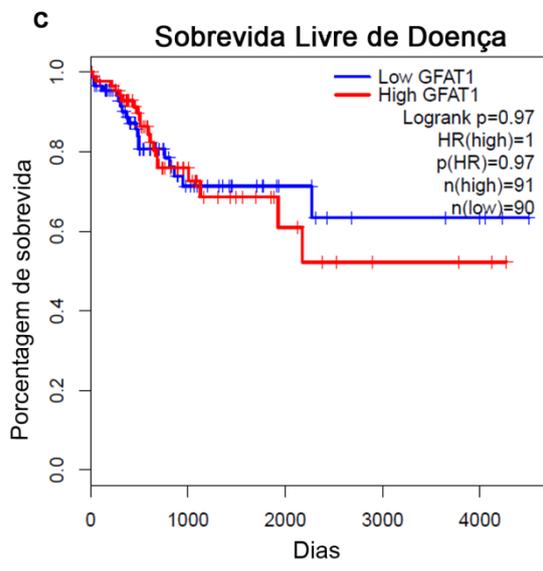
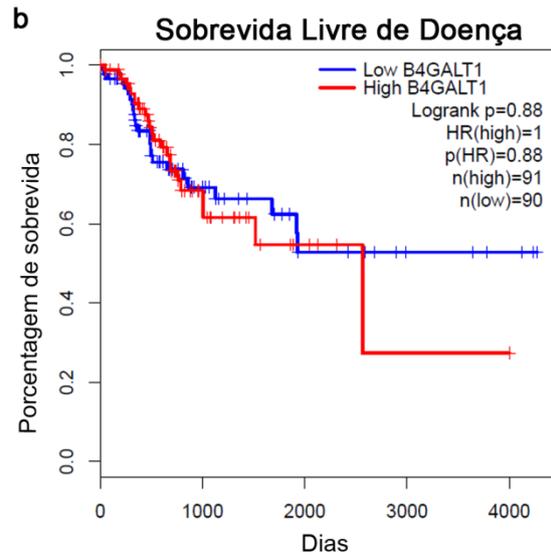
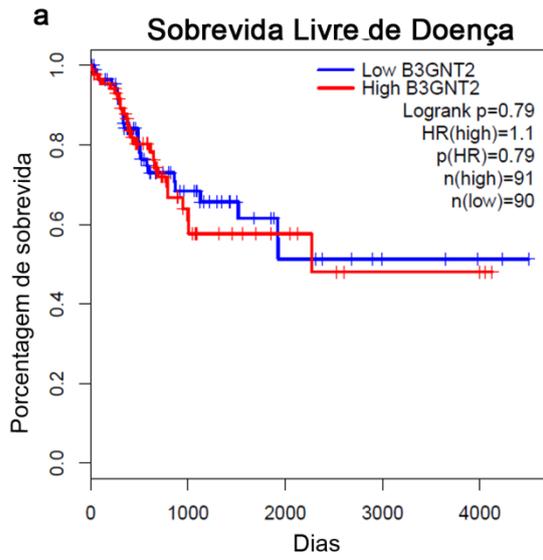


Figura 5.26: Sobrevida global de acordo com os níveis de expressão de diferentes glicogenes em amostras de câncer de cólon. Amostras depositadas no banco de dados COAD do TCGA foram separadas em tercís de expressão e aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevida global dos pacientes. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.



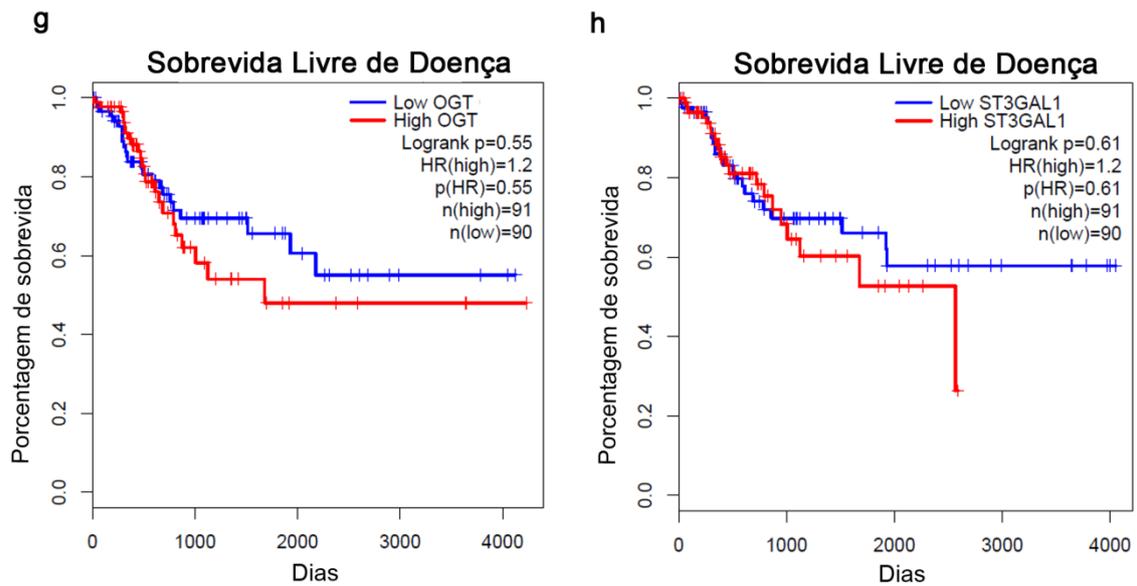


Figura 5.27: Sobrevida livre de doença de acordo com os níveis de expressão de diferentes glicogênes em amostras de câncer de cólon. Amostras depositadas no banco de dados COAD do TCGA foram separadas em tercís de expressão e aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevida livre de doença dos pacientes. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.

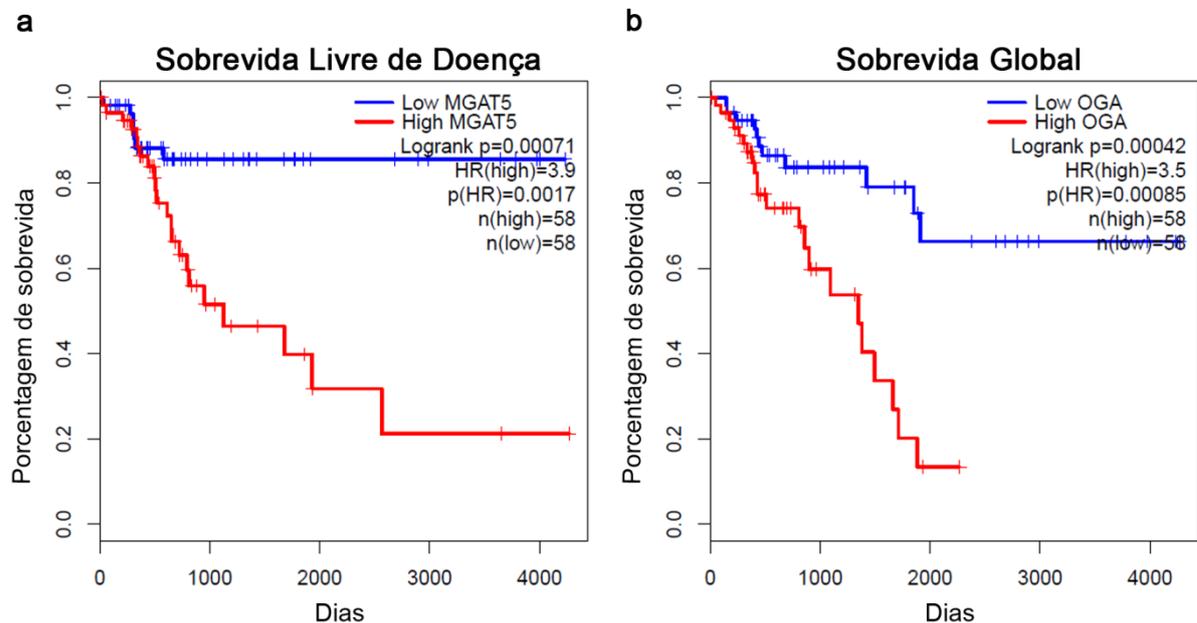


Figura 5.28: Sobrevida livre de doença e sobrevida global de acordo com os níveis de expressão de *MGAT5* e *OGA* apenas de pacientes que não apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados COAD do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSS foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de *MGAT5* na sobrevida livre de doença (a) e o

impacto da expressão de OGA na sobrevida global (b) de pacientes com câncer de cólon que tem MSS. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.

6. **Discussão**

As células tronco tumorais representam uma subpopulação de células heterogênea e plástica, portadora de capacidade tanto de autorrenovação quanto de geração de células progenitoras que estarão sujeitas ao processo de diferenciação. Esses atributos evidenciam a relevância dessas células na origem e manutenção da massa tumoral com toda a sua heterogeneidade e complexidade (HIRATA *et al.*, 2019). Além disso, essas células dispõem de habilidades que contribuem para a resistência às terapias convencionais (SHIBUE; WEINBERG, 2017; STEINBICHLER *et al.*, 2018). Sendo assim, estudos que apontem formas eficientes de identificá-las são de grande valia, pois possibilitam sua melhor caracterização e, conseqüentemente, podem apontar mecanismos de manutenção próprios dessa população que permitam o desenvolvimento de drogas específicas, as quais beneficiariam o paciente com doença metastática ou recorrente. Nesse sentido, o presente estudo se propôs a identificar e caracterizar glicobiomarcadores presentes na população de CTTs. Para isso, primeiramente era preciso empregar uma metodologia que promovesse enriquecimento dessa população. Então, foi escolhido o ensaio de formação de esferas como plataforma de estudo. Nesse ensaio, células em densidade clonal (densidade que permite a geração clonal de esferas) são desafiadas a crescer em condição de não aderência, em presença de meio que favorece o enriquecimento de CTTs e não apresenta indutores de diferenciação. Sabendo que o sucesso desse ensaio depende de parâmetros críticos como a composição do meio, a densidade de células e o tempo de exposição ao meio (PASTRANA; SILVA-VARGAS; DOETSCH, 2011), o presente trabalho se propôs a padronizar um protocolo para o ensaio de formação de esferas para as linhagens de adenocarcinoma de cólon Caco-2 e HT-29. Embora tenha sido possível encontrar uma condição que viabilizava a geração de esferas, as mesmas apresentavam diferentes tamanhos, o que representa uma limitação experimental que pode acarretar em maior variabilidade quando são feitas comparações entre experimentos independentes. Além disso, a quantidade de material gerado por esse ensaio é pequena, inviabilizando a análise das esferas por meio de técnicas que demandem grande quantidade de células.

Uma vez padronizado o protocolo de formação de esferas, era preciso garantir que a população obtida realmente se tratava de CTTs. Para isso, a expressão dos marcadores de CTTs *CD44*, *CD24*, *PROM1*, *ALDH1* e *LGR5* foi

avaliada por RT-qPCR. No entanto, levando em consideração que células crescidas em condições de aderência (2D) e sob a forma de esferas (3D) apresentam consideráveis mudanças morfológicas, antes de iniciar os testes para a expressão dos marcadores foi necessário investigar se isso afetava a expressão de genes-referência. Então, dez genes-referência foram avaliados usando a linhagem Caco-2. Curiosamente, genes-referência amplamente usados como *ACTB* e *18S* apresentaram as maiores diferenças de Cq observadas entre a condição parental e a condição esfera. Já os que apresentaram as menores diferenças de Cq foram *UBE2D2*, *PGK1* e *PPIA*. No entanto, dentre eles, *UBE2D2* e *PPIA* foram mais amplamente expressos em Caco-2 e, por esse motivo, foram os escolhidos. No caso da linhagem HT-29, *UBE2D2* não apresentou amplificação, o que obrigou o uso de apenas *PPIA* como gene-referência, configurando uma limitação do presente estudo. De fato, um grupo que se propôs a identificar os melhores genes-referência para análises de expressão gênica em CTTs também observou grande diferença na expressão de *ACTB*, o que levou os autores a desaconselharem o seu uso. Além disso, entre os genes-referência apontados como mais estáveis, o mesmo estudo também menciona *PPIA* como uma possibilidade (LEMMA *et al.*, 2016).

Tendo estabelecido os genes-referência, foi possível prosseguir com as análises de expressão gênica dos marcadores de CTTs (*CD44*, *CD24*, *PROM1*, *ALDH1* e *LGR5*) em Caco-2 e HT-29. Ao se comparar a condição parental com a condição de esfera, verificou-se que, em Caco-2, houve redução significativa na expressão de *CD24*, e em HT-29 ocorreu aumento significativo na expressão de *ALDH1*, *LGR5* e *PROM1*. Embora ambas as linhagens tenham sido submetidas ao mesmo meio de enriquecimento de CTTs, a diferença encontrada na expressão dos marcadores entre elas evidencia a heterogeneidade da população de CTTs. Curiosamente, essa heterogeneidade foi retratada por um estudo que mostrou existirem diferenças na presença de marcadores de CTTs entre cinco linhagens de CCR. Neste caso, a população CD133⁺ de cada uma das linhagens analisadas foi isolada e outros marcadores relacionados ao perfil tronco, tumorigênese e potencial metastático foram avaliados. Assim, três conjuntos de marcadores foram relatados: I) os de detecção homogênea (*CD9*, *CD29*, *CD49b*, *CD59*, *CD151*, *CD326*), que pareciam representar uma assinatura da população de CTTs devido à sua detecção em todas as linhagens estudadas; II) os de detecção heterogênea (*CD44*, *CD47*, *CD49f*, *CD55*, *CD221*, *CD227*, *CD324*), com diferenças na porcentagem de células positivas entre as linhagens; e III) aqueles de marcação descontínua (*CD24*, *CD26*,

CD54, CD66c, CD81, CD90, CD99, CD112, CD164, CD166, CD200), ora presentes, ora ausentes, dependendo da linhagem celular avaliada (GEMEI *et al.*, 2013). Dessa forma, as diferenças observadas no perfil de expressão dos marcadores de CTTs em Caco-2 e HT-29 podem ser um reflexo da presença de subpopulações com perfis distintos nos esferoides formados.

Em um primeiro momento, a redução na expressão de *CD24* em Caco-2 parece contraditória, uma vez que o aumento desse marcador já foi documentado em populações de CTTs oriundas de linhagens de CCR (KE *et al.*, 2012). Foi sugerido também que *CD24* atua como principal molécula responsável pela formação de *clusters* de CTTs, os quais poderiam ser importantes para a sobrevivência dessas células durante sua disseminação na corrente sanguínea (KAPELERIS *et al.*, 2019). No entanto, essa redução na expressão de *CD24* poderia indicar a presença de uma população de CTTs que apresenta mecanismos de resistência a drogas e de sobrevivência mediados por autofagia, já que o silenciamento de *CD24* em células que superexpressavam este marcador promoveu aumento da autofagia em células de CCR (ZHUO; WANG, 2019). Neste sentido, foi demonstrado também que a inibição do processo autofágico se mostrou capaz de sensibilizar células de CCR ao 5-Fluorouracil (5-FU) e a drogas antiangiogênicas (LI *et al.*, 2009; SELVAKUMARAN *et al.*, 2013). Corroborando estes achados, foi descrito em CCR que o fenótipo $CD133^+/CD24^{low}$ está associado a resistência ao 5-FU (PASCHALL *et al.*, 2016). Assim, a expressão ou não de *CD24* parece conferir vantagens em diferentes processos: metástase ou resistência. Uma hipótese possível é que a exposição da linhagem Caco-2 a um cultivo em suspensão represente um insulto à sua sobrevivência, o que obrigaria essas células a ativarem autofagia através de um mecanismo que envolva a redução na expressão de *CD24*.

No que diz respeito às variações de expressão quando compara-se a condição parental com a condição de esfera nas células HT-29, foi visto um aumento de três marcadores de CTTs, nomeadamente *LGR5*, *PROM1* e *ALDH1*. Este achado é corroborado por dados da literatura, já que todos funcionam como marcadores de CTTs em modelos de CCR (VERMEULEN *et al.*, 2008; HUANG *et al.*, 2009; SHIMOKAWA *et al.*, 2017; ZHANG *et al.*, 2019). *LGR5* foi primeiramente caracterizado como um marcador da população de células tronco intestinais presente na base das criptas, mas tem sido empregado também como marcador de CTTs (KUMAR; BURGESS; GULBIS, 2014; DAME *et al.*, 2018). A proteína *CD133*, codificada pelo gene *PROM1*, em condições não patológicas, está associada à

regiões de protrusão de membrana e microvilos (GLUMAC; LEBEAU, 2018). Um estudo mostrou também que CD133 promove a sinalização de WNT, por meio de um mecanismo de estabilização de β -catenina mediado por HDAC6 (do inglês *Histone deacetylase 6*) e, neste mesmo trabalho, foi observado que o silenciamento de CD133 leva a um aumento da expressão de genes relacionados à diferenciação (MAK *et al.*, 2012). Além disso, também foi demonstrado em um modelo de glioblastoma que a capacidade de autorrenovação e tumorigênese de CTTs era dependente da fosforilação de CD133 (na tirosina 828), a qual mediava sua interação com p85 e promovia a sinalização via PI3K/AKT (WEI *et al.*, 2013). Com relação ao ALDH1, este representa uma das 11 famílias de isoenzimas que catalisam a oxidação de aldeídos a ácidos carboxílicos (VASSALLI, 2019). Populações de CTTs apresentam maior atividade de ALDH e, por isso, essa característica é empregada como o racional do ensaio de Aldefluor® para a identificação de CTTs (MELE; LICCARDO; TIRINO, 2018). Embora no presente estudo a atividade dessa enzima não tenha sido investigada, o aumento da expressão gênica de *ALDH1* dá indícios de que sua atividade pode estar aumentada nessa população. Nesse contexto, já foi demonstrado que a superexpressão independente de nove isoformas de *ALDH* resulta em aumento da população de ALDH+ no ensaio de atividade (ZHOU *et al.*, 2019). Portanto, diante do exposto acima, o aumento da expressão dos marcadores *ALDH1*, *LGR5* e *PROM1* em células HT-29 submetidas ao ensaio de formação de esferas sugere que esteja ocorrendo um enriquecimento da população de CTTs.

No contexto do câncer os glicanos desempenham função em diferentes aspectos da transformação maligna, tais como: proliferação, migração, invasão, escape da vigilância imunológica e metástase (RODRIGUES *et al.*, 2018). No entanto, esses processos estão relacionados a alterações nos níveis de arranjos específicos, tais como: aumento de *N*-glicanos com ramificações do tipo β 1,6-GlcNAc que são geradas pela enzima MGAT5 (Figura 1.9a); aumento de glicanos fucosilados e sialilados; e ainda a presença aumentada de antígenos de carboidratos associados ao tumor, sejam eles epítomos Lewis (Figura 1.9c) ou produtos de *O*-glicanos truncados, como os antígenos T, sT, Tn e sTn (Figura 1.9 b).

Embora CD44, CD24 e CD133 sejam utilizados como marcadores de CTTs, essas proteínas também participam de mecanismos celulares que podem ser impactados mediante alterações na glicosilação. CD44 é uma glicoproteína de membrana envolvida com mecanismos de adesão celular e transdução de sinal que

estão intimamente relacionados à diferenciação, proliferação, migração e sobrevivência (ORIAN-ROUSSEAU; SLEEMAN, 2014). Uma das principais funções de CD44 está relacionada à sua ligação ao ácido hialurônico (KAROUSOU *et al.*, 2017). Já foi demonstrado, em modelos de câncer gástrico, que a modificação de CD44 por O-glicanos truncados é capaz de aumentar a sua interação com o ácido hialurônico (MEREITER *et al.*, 2019). Por outro lado, a presença de ácidos siálicos terminais em N-glicanos de CD44 atuaria no sentido contrário, impedindo essa interação (FALLER; GUVENCH, 2014).

A glicoproteína CD24 consiste numa molécula de adesão a qual é ancorada na membrana em regiões de *lipid rafts* por uma porção GPI (EYVAZI *et al.*, 2018). A função adesiva de CD24 à P-selectina é mediada pela presença do arranjo sLe^x (Figura 1.9c) em CD24 (FRIEDERICHS *et al.*, 2000). Essa interação com a P-selectina presente em células do endotélio vascular possibilita o rolamento das células tumorais na corrente sanguínea, favorecendo o desenvolvimento de metástases (KÖHLER *et al.*, 2010). Além disso, um estudo demonstrou que os N-glicanos de CD24 apresentavam *core* fucosilação (Figura 1.9a), enquanto seus O-glicanos apresentavam forte presença de estruturas sialil T (Figura 1.9b) (MOTARI *et al.*, 2009).

Para a glicoproteína CD133, a perda induzida de um de seus sítios de N-glicosilação impede sua interação com a β -catenina o que resulta em inativação da via de WNT (LIU *et al.*, 2015). Além disso, a presença de sialilação (α -2,3) é capaz de promover redução da cinética de degradação de CD133 (ZHOU *et al.*, 2010).

No presente estudo, os esferoides gerados tiveram a expressão de glicogenes avaliada. Embora não tenha sido possível documentar mudanças estatisticamente significativas na expressão desses genes em Caco-2, foi observado aumento estatisticamente significativo na expressão de *B3GNT2*, *OGT* e *OGA* em HT-29. De forma pioneira em estudos de Glicobiologia do câncer, foram realizadas análises *in silico* que apontam para a existência de correlação positiva entre a expressão de marcadores de CTTs e glicogenes, como é o caso de *CD44* x *OGA*, *CD44* x *B3GNT2*, *LGR5* x *OGA* para amostras de câncer de cólon, e *CD44* x *OGA*, *LGR5* x *OGA*, *CD44* x *OGT*, *CD24* x *B3GNT2* para amostras de câncer de reto. As correlações entre *CD44* x *OGA* e *LGR5* x *OGA* estão presentes independentemente do local do tumor, no entanto, no caso de *CD44* x *B3GNT2*, parece ser válida somente para carcinomas de cólon, enquanto as correlações entre *CD44* x *OGT* e *CD24* x *B3GNT2* foram observadas apenas em carcinomas de reto. Todas essas

correlações descritas foram estatisticamente significativas, no entanto, se mostraram correlações fracas (r entre 0.20 e 0.39) ou muito fracas (r entre 0.00 e 0.19) (EVANS, 1996). Uma possível explicação consiste no fato das análises envolverem subpopulações de CTTs que estão presentes em menor número dentro da massa tumoral. Assim, seria mais difícil encontrar expressão desses marcadores em amostras tumorais, enfraquecendo qualquer correlação que os envolva. Vale ressaltar que a correlação positiva entre a expressão de *LGR5* e *OGA* aponta para uma possível regulação entre esses dois genes, pois, conforme demonstrado no presente estudo, o ensaio de formação de esferas promove aumento na expressão de ambos em HT-29 (Figura 5.9 e 5.10b). Embora seja possível especular que esses glicogenes possam exercer influência sobre o estado tronco e a expressão de marcadores em câncer colorretal, os eventuais mecanismos envolvidos ainda não são conhecidos, evidenciando que diferentes vertentes de investigação serão necessárias para esclarecer este ponto.

Em conformidade com a correlação positiva entre os níveis de expressão de *CD44* e *OGT* observada aqui, recentemente foi demonstrado que em esferas oriundas de linhagens de câncer de mama enriquecidas para CTTs há aumento concomitante de *OGT* e da modificação de proteínas por O-GlcNAc. Ademais, a inibição de *OGT* teve um impacto negativo no estado tronco dessas células, enquanto sua superexpressão teve efeito contrário, pois foi capaz de regular positivamente *CD44*, *Nanog*, *c-Myc* e *KLF8*, indicando que o aumento de *OGT* nas esferas pode ser necessário para a regulação de fatores transcricionais relacionados com a pluripotência (AKELLA *et al.*, 2020).

Embora *OGT* esteja atraindo bastante atenção dos pesquisadores nos últimos anos, por ser a responsável pela transferência de GlcNAc para proteínas citosólicas e nucleares, os achados *in vitro* e *in silico* aqui descritos, que apontam aumento da expressão de *OGA* em câncer colorretal, bem como sua correlação com a expressão de *LGR5* e *CD44*, dão indícios da importância desta outra enzima para as CTTs. Nesse sentido, em camundongos mutados para o gene *APC* e heterozigotos para *Oga* (*Oga*^{+/-}), os quais apresentavam menores níveis de *OGA* e *OGT*, foi observado abrandamento em número e volume de adenomas colorretais, devido à redução na expressão de genes-alvo da via de WNT de maneira independente de β -catenina (YANG *et al.*, 2014). Assim, é possível que o aumento de *OGA* observado nas esferas ocorra para induzir a expressão de genes-alvo da via WNT, que está relacionada à manutenção das células tronco intestinais (KRISHNAMURTHY;

KURZROCK, 2018). Além disso, CD44 e LGR5 são genes-alvo dessa via (WIELENGA *et al.*, 1999; BARKER *et al.*, 2007; ZEILSTRA *et al.*, 2014), o que corrobora a correlação positiva encontrada entre a expressão deles com OGA.

Outro aspecto interessante diz respeito ao aumento concomitante na expressão de OGT e OGA na condição de esferas. Em desacordo com essa observação, pesquisadores mostraram existir um mecanismo compensatório entre OGT e OGA que visa a manutenção do estado de equilíbrio. Para isso, utilizaram um modelo de CCR submetido a desequilíbrio na homeostase do processo de modificação por O-GlcNAc por privação de glicose ou influxo de glutamina. No primeiro cenário, a privação de glicose levou ao aumento de OGT e redução de OGA, ambos em nível proteico e de expressão gênica, enquanto no segundo cenário, o influxo de glutamina levou à redução apenas dos níveis proteicos de OGT e aumento dos níveis de OGA e de seu mRNA, apontando para uma regulação em nível transcricional para OGA e pós transcricional para OGT (DECOURCELLE *et al.*, 2020). Curiosamente, esse mesmo trabalho também mostrou que o *knockdown* de OGT reduzia os níveis de OGA e de seu mRNA, do mesmo modo que o contrário também ocorria, contudo de forma mais branda. Esse dado curioso também foi observado em hepatócitos de camundongo, no entanto, neste caso foi demonstrado que embora a superexpressão de Oga cause aumento de Ogt, o contrário não ocorre. De fato, OGA parece regular a transcrição de OGT através de um mecanismo de cooperação entre OGA e p300 (uma histona acetiltransferase que funciona como coativador da transcrição) para a transcrição de OGT (QIAN *et al.*, 2018). Se o mesmo ocorrer em células de CCR, é provável que a expressão aumentada de OGA nas esferas induza também o aumento dos níveis de expressão de OGT, porém o papel que esse aumento concomitante teria em CTTs ainda precisa ser melhor explorado. Uma das hipóteses seria que OGT e OGA se comportassem como os marcadores epigenéticos observados em células tronco quiescentes. Nesse contexto, os genes apresentam perfil bivalente, com presença concomitante de ambos os marcadores repressivos e ativadores, como forma de estarem preparados para qualquer um dos cenários: inativação ou ativação (LIU *et al.*, 2013). Assim, a presença de níveis aumentados de ambas as enzimas poderia representar um mecanismo das células tronco tumorais para que estejam preparadas tanto pra permanecerem no estado tronco ou darem origem a uma célula progenitora.

No presente trabalho foi verificado também que os níveis de expressão de *B3GNT2* correlacionaram-se positivamente com os de *CD44* (câncer de cólon) e *CD24* (câncer de reto). A enzima β 3GNT2 (codificada pelo gene *B3GNT2*), junto com β 4GALT1 (codificada pelo gene *B4GALT1*), são as principais envolvidas na síntese de LacNAc. Essas estruturas, uma vez formadas, podem ser decoradas com antígenos Lewis (TOGAYACHI *et al.*, 2010; BYDLINSKI *et al.*, 2018) (Figura 1.9c). Além disso, a presença de estruturas poli-LacNAc (formadas pela associação de unidades LacNAc) em receptores de membrana possibilita sua interação com galectinas (proteínas com domínios do tipo lectina que se ligam a carboidratos), favorecendo assim a estabilização deles na membrana, o que prolonga a duração da sinalização celular. Um grupo que caracterizou *N*-glicanos presentes na forma padrão de *CD44* (*CD44s*, originada a partir da junção dos éxons 1-5 e 16-20) ressaltou a presença predominante de *N*-glicanos complexos com reduzida presença de fucose e ácido siálico, inclusive no sítio N5, o qual é apontado como o responsável pela ligação ao ácido hialurônico presente na matriz extracelular (HAN *et al.*, 2012). Essa menor decoração dos *N*-glicanos de *CD44* com ácido siálico, pelo menos no sítio N5, pode ter relação com sua interação com o ácido hialurônico, já que a presença de sialilação α -2,3 em *CD44* foi apontada como motivo de redução da interação entre essas duas moléculas (FALLER; GUVENCH, 2014). Conjuntamente, estes achados mostram que a correlação encontrada entre a expressão de *CD44* x *B3GNT2* necessita ser melhor estudada.

Com relação à presença de antígenos Lewis na extremidade de poli-LacNAc, já foi relatado que o marcador *CD24* pode apresentar o epítipo sLe^x (Figura 1.9c), que estaria relacionado com sua capacidade de adesão ao endotélio vascular por meio de interação com P-selectinas (FRIEDERICHS *et al.*, 2000), favorecendo o processo de metástase. Esse fato possibilita levantar uma hipótese de conexão mecânica entre *CD24* e estruturas poli-LacNAc, auxiliando assim a compreensão da correlação vista entre a expressão de *CD24* x *B3GNT2* em amostras de carcinoma retal.

Recentemente, uma investigação revelou a presença de características tronco em células tumorais circulantes (CTC) de pacientes com CCR metastático (GRILLET *et al.*, 2017), este achado motivou a avaliação do impacto clínico da expressão de marcadores de CTTs e de glicogenes em pacientes com CCR. Assim, foram realizadas análises *in silico* com amostras oriundas do TCGA. Foi possível observar que amostras de carcinoma de cólon apresentavam aumento na expressão de

LGR5, *CD44* e redução na expressão de *CD24* e *ALDH1*, quando comparadas com amostras de tecido adjacente não tumoral. O mesmo padrão de expressão de *LGR5*, *CD44* e *ALDH1* ocorreu para as amostras de carcinoma de reto, no entanto, nessas, também se observou aumento na expressão de *PROM1*. Embora não tenha separado as análises em carcinomas do cólon e reto, por meio de imunohistoquímica de amostras normais e de câncer colorretal, um grupo confirmou, pelo menos em parte, os achados *in silico* aqui apresentados. Neste caso, foi verificado um aumento de *CD133* e da forma padrão de *CD44*, além de uma tênue redução de *ALDH1* nas amostras de tumor primário, em comparação feita com amostras normais (LUGLI *et al.*, 2010). Outro estudo também contribuiu com os achados aqui demonstrados, ao apontar o aumento da marcação de *LGR5* em amostras de tecidos de CCR pareadas (HE *et al.*, 2014).

Na avaliação dos marcadores de CTTs nos diferentes estádios do câncer de cólon, verificou-se que a expressão de *LGR5* e *CD44* está aumentada em todos os estágios da progressão da doença. Enquanto o contrário ocorreu para *CD24*. Já para *ALDH1*, a redução na expressão alcançou significância estatística a partir do estadio II, enquanto a expressão de *CD44* teve uma redução estatisticamente significativa do estadio I para o estadio IV, o que poderia indicar que para o processo de metástase distante seria necessária a redução na expressão de *CD44*. Já para amostras de carcinoma de reto, o mesmo cenário se repete para as expressões de *LGR5* e *CD44*, que se mantêm aumentadas ao longo de todos os estádios. Por outro lado, o aumento na expressão de *PROM1* só ocorre do estadio I ao estadio III, o que pode indicar que *PROM1* exerce algum papel na progressão da doença, mas não para o processo de metástase distante. Além disso, em câncer de reto há redução na expressão de *ALDH1* também ao longo de todos os estádios. No entanto, nenhum dos marcadores analisados foi capaz de prever a sobrevida livre de doença ou a sobrevida global. Dentre todos os marcadores analisados, apenas a expressão de *ALDH1*, em pacientes com câncer de reto, parece ter potencial de prever a sobrevida global, uma vez que o *P* valor do *logrank* foi de 0,056, porém, devido à quantidade reduzida de amostras (apenas 31 no grupo com alta expressão e 32 no grupo com baixa expressão), não foi possível atingir a significância estatística.

Pela análise de imunohistoquímica de 523 amostras de tecido de adenocarcinoma colorretal, um trabalho relacionou os níveis de *CD133*, *CD44* e *CD24* com características clinicopatológicas, e concluiu que *CD133* estava

relacionado aos estadios T (que diz respeito à extensão do tumor primário) mais avançados, CD24 estava relacionado ao grau de diferenciação e CD44 ao tamanho do tumor. No entanto, nenhum deles teve impacto na sobrevida global de pacientes com CCR (CHOI *et al.*, 2009), de forma análoga ao que foi apontado pelas análises *in silico* aqui apresentadas.

Outro aspecto interessante diz respeito a localização subcelular/tecidual de ALDH. Um estudo que avaliou por imunohistoquímica os níveis proteicos de ALDH1A1 (uma das 19 isoenzimas da família de aldeído desidrogenases) em amostras de tecido de câncer de cólon e reto encontrou diferentes tipos de marcação: citoplasmática, nuclear e estromal. Curiosamente, apenas a presença de ALDH1A1 nuclear resultava em menor sobrevida global para pacientes com câncer de cólon (KAHLERT *et al.*, 2012). Outro estudo apontou que amostras de câncer colorretal avaliadas para ALDH1A1 pela mesma técnica anterior tinham maior marcação em comparação com as amostras de tecido normal adjacente, e isso estava associado à presença de metástases linfonodais (YANG *et al.*, 2018). No entanto, amostras de metástase hepática apresentaram maiores níveis de ALDH1A1 que as amostras pareadas de seus tumores primários, porém, em qualquer um dos cenários, seus níveis eram aumentados se os pacientes tivessem sido submetidos a tratamento antes da ressecção. Além disso, seus níveis foram associados a tumores pouco diferenciados, localizados do lado direito do cólon (VAN DER WAALS; BOREL RINKES; KRANENBURG, 2018). Assim, uma hipótese pode ser formulada para explicar as incongruências encontradas. É possível que as amostras analisadas nessa dissertação fossem em sua maioria de tumores localizados do lado esquerdo, já que o estudo citado menciona a diminuição nos níveis de ALDH1A1 em cólon descendente, reto e sigmoide.

As mesmas análises *in silico* realizadas para os marcadores de CTTs foram repetidas para os glicogenes. Em amostras de carcinoma de cólon foi observado aumento da expressão de *MGAT5*, *OGA*, *OGT*, *ST3GAL1* e redução da expressão de *B4GALT1*, em relação às amostras de tecido não tumoral adjacente. Por outro lado, em amostras de carcinoma de reto, ocorreu redução na expressão de *B3GNT2* e *ST3GAL1* e aumento de *OGT*.

Reforçando os achados encontrados por bioinformática em cólon, um trabalho mostrou o aumento da expressão de *MGAT5* em amostras de adenomas e carcinomas colorretais quando estas são comparadas com o tecido normal adjacente (PETRETTI *et al.*, 2000). Os *N*-glicanos com ramificações do tipo β 1,6-

GlcNAc, que são gerados por essa enzima, também foram mais detectados por imunohistoquímica em amostras de carcinoma colorretal do que em seus pares não tumorais (HÄGERBÄUMER *et al.*, 2015).

No que tange a *ST3GAL1*, dois trabalhos distintos evidenciaram o aumento de sua expressão em amostras de carcinoma colorretal quando comparadas com o tecido adjacente não tumoral (PETRETTI *et al.*, 2000; SCHNEIDER *et al.*, 2001). Todavia, nossas análises *in silico* demonstraram antagonismo na expressão desse glicogene entre amostras de carcinoma de cólon e reto, o que pode ter relação com a lateralidade. Embora o CCR seja tratado como uma única entidade, a origem embrionária do cólon proximal é a partir do intestino médio primitivo, enquanto o cólon distal e reto é a partir do intestino posterior primitivo. Além disso, os tumores proximais de cólon estão mais associados à mutações em *BRAF* e ao CIMP, enquanto carcinomas do cólon distal e reto estão mais associados a mutações em *APC* e *TP53* e à CIN (YOUNG; GOLZARIAN, 2018). Curiosamente, análises de *microarray* de amostras de carcinoma de cólon e reto apontaram maior desregulação de vias de sinalização envolvidas com processos metabólicos e biossíntese de *N*- e *O*-glicanos nos tecidos de reto (LI *et al.*, 2012).

Um grupo que analisou os níveis de OGT, OGA e de modificação por *O*-GlcNAc em amostras pareadas de pacientes com adenocarcinoma de cólon, apontou um aumento tanto de OGT quanto da modificação por *O*-GlcNAc e uma tendência de aumento de OGA nos tumores colorretais, em comparação com amostras adjacentes não tumorais (PHUEAOUAN *et al.*, 2013).

Para além das análises comparativas entre amostras pareadas, também foram realizadas aqui avaliações da expressão dos glicogenes por estadiamento. No carcinoma de cólon, foi possível perceber que a expressão de *MGAT5* e *OGT* é aumentada ao longo de toda a progressão da doença, o que, nesse último caso, também se repete para carcinoma de reto. Os níveis de expressão de *OGT* aumentam do estadio II para o IV, o que sugere um papel de *OGT* no estabelecimento de metástases distantes. De fato, a maior expressão de *OGT* foi correlacionada positivamente com o estadiamento total e com estadios N e M. Além disso, maiores níveis de *OGT* reduziram a sobrevida global em três anos de pacientes com CCR (XU *et al.*, 2019). No entanto, as análises envolvendo o desfecho clínico feitas pelo presente estudo não demonstraram tal redução.

Com relação à *MGAT5*, níveis aumentados dessa enzima (também conhecida como GnT-V) foram correlacionados com metástases distantes e com menor

sobrevida global em pacientes com CCR (MURATA *et al.*, 2000). No entanto, os autores ressaltam que essa menor sobrevida global refletiu, em parte, a grande quantidade de pacientes estadio IV e a pior sobrevida dos pacientes em estadio II que apresentavam maiores níveis de *MGAT5*, o que pode explicar a discrepância com as análises aqui apresentadas. Porém, no que diz respeito à sobrevida livre de doença, a maior expressão de *MGAT5* foi aqui apontada como pior prognóstico.

No presente trabalho, também foi visto que a expressão de *OGA* tem aumento significativo apenas nos estadios tardios III e IV. Porém, estudos *in vitro* utilizando as células de tumor primário (SW480) e de metástase linfonodal (SW620) mostraram que nessa última linhagem havia aumento da modificação por O-GlcNAc e redução na expressão de *OGA*. Uma vez que SW620 era silenciada para *OGA* passava a apresentar características relacionadas ao processo de EMT com aquisição de morfologia fibroblastoide e alteração na expressão de diversos genes envolvidos com processos de motilidade (YEHEZKEL *et al.*, 2012). No entanto, na contramão desses dados, os resultados aqui apresentados demonstram que a maior expressão de *OGA* foi relacionada a uma pior sobrevida global em pacientes com carcinoma de cólon.

As análises para carcinoma retal apontaram pequena redução da expressão de *ST3GAL1* no estadio III, com posterior aumento discreto no estadio IV, o que sugere um possível papel de *ST3GAL1* no desenvolvimento de metástases distantes. Vale ressaltar que em câncer de mama, a alta expressão de *ST3GAL1* em pacientes com estadiamento avançado foi associado à menor sobrevida livre de recorrência (FAN *et al.*, 2018). É possível que a não estratificação dos pacientes por estadiamento antes das análises de sobrevida tenha inviabilizado essa mesma observação para câncer colorretal.

Sob a perspectiva das CTTs, fica evidente que determinadas alterações que ocorrem no glicoma dessas células podem ter impacto na sobrevida de pacientes com CCR. Porém, é importante ressaltar que todas as alterações observadas no presente estudo foram investigadas em relação à expressão gênica. Essas alterações não necessariamente refletem alterações em nível proteico, tampouco refletem maior ou menor presença de determinado arranjo em glicanos. Dessa forma, análises complementares – dos níveis proteicos dos marcadores de CTTs, das enzimas codificadas pelos glicogenes, bem como dos arranjos gerados por elas – necessitam ser realizadas. Com relação aos dados *in silico*, a incapacidade de separação das amostras do TCGA de acordo com os subtipos moleculares CMS e

CRIS em ambas as plataformas utilizadas (GEPIA e UALCAN) impediu o estabelecimento de associação entre determinado subtipo molecular e a maior presença de população com perfil tronco. Assim, o papel exato que cada um desses glicogenes e suas respectivas enzimas desempenham na regulação do estado tronco ainda precisa ser desvendado.

7. Conclusão

O ensaio de formação de esferas aplicado a linhagens de CCR com diferentes perfis mutacionais e de diferenciação promove enriquecimento de subpopulações de CTTs que expressam marcadores de superfície distintos. Isso indica que mesmo em condições em que as células são expostas a um meio de enriquecimento de CTTs contendo suplementação idêntica, cada linhagem responde de maneira particular a esses estímulos. Além disso, a alta plasticidade das CTTs torna essa metodologia bastante variável, pois existe importante variabilidade na expressão dos marcadores em cada réplica biológica. Ainda assim, os achados *in vitro* evidenciaram a presença de alterações nos glicogenes *OGT*, *OGA* e *B3GNT2* em populações de CTTs com maior expressão dos marcadores *ALDH1*, *LGR5* e *PROM1*. Mais do que apenas expressão alterada, esses glicogenes apresentaram correlação positiva com a expressão de determinados marcadores, e um deles (*OGA*) foi capaz de distinguir um grupo de pacientes com câncer de cólon que apresentava menor sobrevida global, o que, pelo menos nesse trabalho, não foi viável para nenhum dos marcadores de CTTs. Portanto, embora representem ainda contribuições modestas, os resultados obtidos com essa pesquisa apontam para um possível papel da glicosilação aberrante na manutenção do fenótipo tronco e abrem caminho para que, posteriormente, algum glicobiomarcador possa ser incorporado como marcador da população de CTTs.

8. Referências bibliográficas

AHMED, D. *et al.* Epigenetic and genetic features of 24 colon cancer cell lines. **Oncogenesis**, 2, p. e71, Set 2013.

AKBARZADEH, M. *et al.* Current approaches in identification and isolation of cancer stem cells. **J Cell Physiol**, Fev 2019.

AKELLA, N. M. *et al.* O-GlcNAc Transferase Regulates Cancer Stem-like Potential of Breast Cancer Cells. **Mol Cancer Res**, Jan 2020.

AL-HAJJ, M. *et al.* Prospective identification of tumorigenic breast cancer cells. **Proc Natl Acad Sci U S A**, 100, n. 7, p. 3983-3988, Abr 2003.

ALDERDICE, M. *et al.* Prospective patient stratification into robust cancer-cell intrinsic subtypes from colorectal cancer biopsies. **J Pathol**, 245, n. 1, p. 19-28, Mai 2018.

ALISSON-SILVA, F. *et al.* Evidences for the involvement of cell surface glycans in stem cell pluripotency and differentiation. **Glycobiology**, 24, n. 5, p. 458-468, Mai 2014.

ARAN, V. *et al.* Colorectal Cancer: Epidemiology, Disease Mechanisms and Interventions to Reduce Onset and Mortality. **Clin Colorectal Cancer**, 15, n. 3, p. 195-203, Set 2016.

ARNOLD, M. *et al.* Global patterns and trends in colorectal cancer incidence and mortality. **Gut**, 66, n. 4, p. 683-691, Abr 2017.

BARKEER, S. *et al.* Novel role of O-glycosyltransferases GALNT3 and B3GNT3 in the self-renewal of pancreatic cancer stem cells. **BMC Cancer**, 18, n. 1, p. 1157, Nov 2018.

BARKER, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. **Nature**, 449, n. 7165, p. 1003-1007, Out 2007.

BATLLE, E.; CLEVERS, H. Cancer stem cells revisited. **Nat Med**, 23, n. 10, p. 1124-1134, Out 2017.

BECKER, J. L.; TRAN, D. T.; TABAK, L. A. Members of the GalNAc-T family of enzymes utilize distinct Golgi localization mechanisms. **Glycobiology**, 28, n. 11, p. 841-848, Nov 2018.

BLEAU, A. M. *et al.* PTEN/PI3K/Akt pathway regulates the side population phenotype and ABCG2 activity in glioma tumor stem-like cells. **Cell Stem Cell**, 4, n. 3, p. 226-235, Mar 2009.

BONNET, D.; DICK, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. **Nat Med**, 3, n. 7, p. 730-737, Jul 1997.

BRAY, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA Cancer J Clin**, 68, n. 6, p. 394-424, Nov 2018.

BYDLINSKI, N. *et al.* The contributions of individual galactosyltransferases to protein specific N-glycan processing in Chinese Hamster Ovary cells. **J Biotechnol**, 282, p. 101-110, Set 2018.

CARETHERS, J. M.; JUNG, B. H. Genetics and Genetic Biomarkers in Sporadic Colorectal Cancer. **Gastroenterology**, 149, n. 5, p. 1177-1190.e1173, Out 2015.

CHANDRASHEKAR, D. S. *et al.* UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. **Neoplasia**, 19, n. 8, p. 649-658, Ago 2017.

CHEN, J. *et al.* A restricted cell population propagates glioblastoma growth after chemotherapy. **Nature**, 488, n. 7412, p. 522-526, 2012.

CHOI, D. *et al.* Cancer stem cell markers CD133 and CD24 correlate with invasiveness and differentiation in colorectal adenocarcinoma. **World J Gastroenterol**, 15, n. 18, p. 2258-2264, Mai 2009.

CUMMINGS, R. D. "Stuck on sugars - how carbohydrates regulate cell adhesion, recognition, and signaling". **Glycoconj J**, 36, n. 4, p. 241-257, Ago 2019.

CUMMINGS, R. D.; PIERCE, J. M. The challenge and promise of glycomics. **Chem Biol**, 21, n. 1, p. 1-15, Jan 2014.

DAGOGO-JACK, I.; SHAW, A. T. Tumour heterogeneity and resistance to cancer therapies. **Nat Rev Clin Oncol**, 15, n. 2, p. 81-94, Fev 2018.

DAME, M. K. *et al.* Identification, isolation and characterization of human LGR5-positive colon adenoma cells. **Development**, 145, n. 6, Mar 2018.

DE PALMA, F. D. E. *et al.* The Molecular Hallmarks of the Serrated Pathway in Colorectal Cancer. **Cancers (Basel)**, 11, n. 7, Jul 2019.

DE SOUSA E MELO, F. *et al.* A distinct role for Lgr5(+) stem cells in primary and metastatic colon cancer. **Nature**, 543, n. 7647, p. 676-680, Mar 2017.

DE' ANGELIS, G. L. *et al.* Microsatellite instability in colorectal cancer. **Acta Biomed**, 89, n. 9-S, p. 97-101, Dez 2018.

DECOURCELLE, A. *et al.* Evidence of a compensatory regulation of colonic O-GlcNAc transferase and O-GlcNAcase expression in response to disruption of O-GlcNAc homeostasis. **Biochem Biophys Res Commun**, 521, n. 1, p. 125-130, Jan 2020.

DESAI, A.; WEBB, B.; GERSON, S. L. CD133+ cells contribute to radioresistance via altered regulation of DNA repair genes in human lung cancer cells. **Radiother Oncol**, 110, n. 3, p. 538-545, Mar 2014.

DIEHN, M. *et al.* Association of reactive oxygen species levels and radioresistance in cancer stem cells. **Nature**, 458, n. 7239, p. 780-783, Abr 2009.

DORSETT, K. A. *et al.* Sox2 promotes expression of the ST6Gal-I glycosyltransferase in ovarian cancer cells. **J Ovarian Res**, 12, n. 1, p. 93, Out 2019.

DREWES, J. L.; HOUSSEAU, F.; SEARS, C. L. Sporadic colorectal cancer: microbial contributors to disease prevention, development and therapy. **Br J Cancer**, 115, n. 3, p. 273-280, Jul 2016.

DRIESSENS, G. *et al.* Defining the mode of tumour growth by clonal analysis. **Nature**, 488, n. 7412, p. 527-530, Ago 2012.

DUNNE, P. D. *et al.* Cancer-cell intrinsic gene expression signatures overcome intratumoural heterogeneity bias in colorectal cancer patient classification. **Nat Commun**, 8, p. 15657, Mai 2017.

DUNNE, P. D. *et al.* Challenging the Cancer Molecular Stratification Dogma: Intratumoral Heterogeneity Undermines Consensus Molecular Subtypes and Potential Diagnostic Value in Colorectal Cancer. **Clin Cancer Res**, 22, n. 16, p. 4095-4104, Ago 2016.

EVANS, J. D. **Straightforward statistics for the behavioral sciences**. Thomson Brooks/Cole Publishing Co, 1996. 0534231004.

EYVAZI, S. *et al.* Involvement of CD24 in Multiple Cancer Related Pathways Makes It an Interesting New Target for Cancer Therapy. **Curr Cancer Drug Targets**, 18, n. 4, p. 328-336, 2018.

FALLER, C. E.; GUVENCH, O. Terminal sialic acids on CD44 N-glycans can block hyaluronan binding by forming competing intramolecular contacts with arginine sidechains. **Proteins**, 82, n. 11, p. 3079-3089, Nov 2014.

FAN, T. C. *et al.* Reciprocal feedback regulation of ST3GAL1 and GFRA1 signaling in breast cancer cells. **Cancer Lett**, 434, p. 184-195, Out 2018.

FEARON, E. R.; VOGELSTEIN, B. A genetic model for colorectal tumorigenesis. **Cell**, 61, n. 5, p. 759-767, Jun 1990.

FERREIRA, I. G. *et al.* Glycosylation as a Main Regulator of Growth and Death Factor Receptors Signaling. **Int J Mol Sci**, 19, n. 2, Fev 2018.

FESSLER, E.; MEDEMA, J. P. Colorectal Cancer Subtypes: Developmental Origin and Microenvironmental Regulation. **Trends Cancer**, 2, n. 9, p. 505-518, Set 2016.

FIDLER, M. M.; SOERJOMATARAM, I.; BRAY, F. A global view on cancer incidence and national levels of the human development index. **Int J Cancer**, 139, n. 11, p. 2436-2446, Dez 2016.

FRIEDERICHS, J. *et al.* The CD24/P-selectin binding pathway initiates lung arrest of human A125 adenocarcinoma cells. **Cancer Res**, 60, n. 23, p. 6714-6722, Dez 2000.

FUENTES-GARCÍA, G. *et al.* O-GlcNAcylation Is Involved in the Regulation of Stem Cell Markers Expression in Colon Cancer Cells. **Front Endocrinol (Lausanne)**, 10, p. 289, Mai 2019.

GEMEI, M. *et al.* Cytometric profiling of CD133+ cells in human colon ^[1]_[SEP]carcinoma cell lines identifies a common core phenotype ^[1]_[SEP]and cell type-specific mosaics. **Int J Biol Markers**, 28, n. 3, p. 267-273, Set 2013.

GIGLIA, M. D.; CHU, D. I. Familial Colorectal Cancer: Understanding the Alphabet Soup. **Clin Colon Rectal Surg**, 29, n. 3, p. 185-195, Set 2016.

GLUMAC, P. M.; LEBEAU, A. M. The role of CD133 in cancer: a concise review. **Clin Transl Med**, 7, n. 1, p. 18, Jul 2018.

GOTO, T.; MARUSAWA, H.; CHIBA, T. Landscape of genetic aberrations detected in human colorectal cancers. **Gastroenterology**, 145, n. 3, p. 686-688, Set 2013.

GRILLET, F. *et al.* Circulating tumour cells from patients with colorectal cancer have cancer stem cell hallmarks in *ex vivo* culture. **Gut**, 66, n. 10, p. 1802-1810, Out 2017.

GUINNEY, J. *et al.* The consensus molecular subtypes of colorectal cancer. **Nat Med**, 21, n. 11, p. 1350-1356, Nov 2015.

GUO, H.; NAGY, T.; PIERCE, M. Post-translational glycoprotein modifications regulate colon cancer stem cells and colon adenoma progression in *Apc*(min/+) mice through altered Wnt receptor signaling. **J Biol Chem**, 289, n. 45, p. 31534-31549, Nov 2014.

GUO, H. *et al.* O-Linked *N*-Acetylglucosamine (O-GlcNAc) Expression Levels Epigenetically Regulate Colon Cancer Tumorigenesis by Affecting the Cancer Stem Cell Compartment via Modulating Expression of Transcriptional Factor *MYBL1*. **J Biol Chem**, 292, n. 10, p. 4123-4137, Mar 2017.

HAN, H. *et al.* Comprehensive characterization of the N-glycosylation status of CD44s by use of multiple mass spectrometry-based techniques. **Anal Bioanal Chem**, 404, n. 2, p. 373-388, Ago 2012.

HANKEY, W.; FRANKEL, W. L.; GRODEN, J. Functions of the APC tumor suppressor protein dependent and independent of canonical WNT signaling: implications for therapeutic targeting. **Cancer Metastasis Rev**, 37, n. 1, p. 159-172, Mar 2018.

HAROSH-DAVIDOVICH, S. B.; KHALAILA, I. O-GlcNAcylation affects β -catenin and E-cadherin expression, cell motility and tumorigenicity of colorectal cancer. **Exp Cell Res**, 364, n. 1, p. 42-49, Mar 2018.

HART, G. W. Nutrient regulation of signaling and transcription. **Journal of Biological Chemistry**, 294, n. 7, p. 2211-2231, Fev 2019.

HE, S. *et al.* Expression of *Lgr5*, a marker of intestinal stem cells, in colorectal cancer and its clinicopathological significance. **Biomed Pharmacother**, 68, n. 5, p. 507-513, Jun 2014.

HIRATA, A. *et al.* Heterogeneity of Colon Cancer Stem Cells. **Adv Exp Med Biol**, 1139, p. 115-126, Mai 2019.

HOLST, S.; WUHRER, M.; ROMBOUTS, Y. Glycosylation characteristics of colorectal cancer. **Adv Cancer Res**, 126, p. 203-256, 2015.

HUANG, E. H. *et al.* Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC overpopulation during colon tumorigenesis. **Cancer Res**, 69, n. 8, p. 3382-3389, Abr 2009.

HUNG, J. S. *et al.* C1GALT1 overexpression promotes the invasive behavior of colon cancer cells through modifying O-glycosylation of FGFR2. **Oncotarget**, 5, n. 8, p. 2096-2106, Abr 2014.

HÄGERBÄUMER, P. Lectin Histochemistry Shows WGA, PHA-L and HPA Binding Increases During Progression of Human Colorectal Cancer. **Anticancer Res**, 35, n. 10, p. 5333-5339, Out 2015.

ISELLA, C. *et al.* Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. **Nat Commun**, 8, p. 15107, Mai 2017.

ISHAQUE, N. *et al.* Whole genome sequencing puts forward hypotheses on metastasis evolution and therapy in colorectal cancer. **Nat Commun**, 9, n. 1, p. 4782, Nov 2018.

ISHIZAWA, K. *et al.* Tumor-initiating cells are rare in many human tumors. **Cell Stem Cell**, 7, n. 3, p. 279-282, Set 2010.

JIANG, Z. *et al.* β 3GnT8 Promotes Colorectal Cancer Cells Invasion via CD147/MMP2/Galectin3 Axis. **Front Physiol**, 9, p. 588, 2018.

KAHLERT, C. *et al.* Expression analysis of aldehyde dehydrogenase 1A1 (ALDH1A1) in colon and rectal cancer in association with prognosis and response to chemotherapy. **Ann Surg Oncol**, 19, n. 13, p. 4193-4201, Dez 2012.

KAPELERIS, J. *et al.* Cancer stemness contributes to cluster formation of colon cancer cells and high metastatic potentials. **Clin Exp Pharmacol Physiol**, Dez 2019.

KAROUSOU, E. *et al.* Roles and targeting of the HAS/hyaluronan/CD44 molecular system in cancer. **Matrix Biol**, 59, p. 3-22, 05 2017.

KE, J. *et al.* A subpopulation of CD24⁺ cells in colon cancer cell lines possess stem cell characteristics. **Neoplasia**, 59, n. 3, p. 282-288, Fev 2012.

KELLER, L.; PANTEL, K. Unravelling tumour heterogeneity by single-cell profiling of circulating tumour cells. **Nat Rev Cancer**, 19, n. 10, p. 553-567, Out 2019.

KIM, Y. S.; KANG, M. J.; CHO, Y. M. Low production of reactive oxygen species and high DNA repair: mechanism of radioresistance of prostate cancer stem cells. **Anticancer Res**, 33, n. 10, p. 4469-4474, Out 2013.

KÖHLER, S. *et al.* E-/P-selectins and colon carcinoma metastasis: first in vivo evidence for their crucial role in a clinically relevant model of spontaneous metastasis formation in the lung. **Br J Cancer**, 102, n. 3, p. 602-609, Fev 2010.

KRISHNAMURTHY, N.; KURZROCK, R. Targeting the Wnt/beta-catenin pathway in cancer: Update on effectors and inhibitors. **Cancer Treat Rev**, 62, p. 50-60, Jan 2018.

KUMAR, K. K.; BURGESS, A. W.; GULBIS, J. M. Structure and function of LGR5: an enigmatic G-protein coupled receptor marking stem cells. **Protein Sci**, 23, n. 5, p. 551-565, Mai 2014.

LAN, X. *et al.* Fate mapping of human glioblastoma reveals an invariant stem cell hierarchy. **Nature**, 549, n. 7671, p. 227-232, Set 2017.

LEMMA, S.; AVNET, S.; SALERNO, M.; CHANO, T. *et al.* Identification and Validation of Housekeeping Genes for Gene Expression Analysis of Cancer Stem Cells. **PLoS One**, 11, n. 2, p. e0149481, Fev 2016.

LI, C. *et al.* Prognostic and clinicopathological value of MUC1 expression in colorectal cancer: A meta-analysis. **Medicine (Baltimore)**, 98, n. 9, p. e14659, Mar 2019.

LI, J. *et al.* Inhibition of autophagy by 3-MA enhances the effect of 5-FU-induced apoptosis in colon cancer cells. **Ann Surg Oncol**, 16, n. 3, p. 761-771, Mar 2009.

LIAU, B. B. *et al.* Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. **Cell Stem Cell**, 20, n. 2, p. 233-246.e237, Fev 2017.

LIN, Y. U. *et al.* LGR5 promotes the proliferation of colorectal cancer cells via the Wnt/ β -catenin signaling pathway. **Oncol Lett**, 9, n. 6, p. 2859-2863, Jun 2015.

LIU, A.; YU, X.; LIU, S. Pluripotency transcription factors and cancer stem cells: small genes make a big difference. **Chin J Cancer**, 32, n. 9, p. 483-487, Set 2013.

LIU, L. *et al.* Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. **Cell Rep**, 4, n. 1, p. 189-204, Jul 2013.

LIU, Y. *et al.* Mutation of N-linked glycosylation at Asn548 in CD133 decreases its ability to promote hepatoma cell growth. **Oncotarget**, 6, n. 24, p. 20650-20660, Ago 2015.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻ ΔΔCT method. **methods**, 25, n. 4, p. 402-408, Dez 2001.

LUGLI, A. *et al.* Prognostic impact of the expression of putative cancer stem cell markers CD133, CD166, CD44s, EpCAM, and ALDH1 in colorectal cancer. **Br J Cancer**, 103, n. 3, p. 382-390, Jul 2010.

MA, H. *et al.* Pathology and genetics of hereditary colorectal cancer. **Pathology**, 50, n. 1, p. 49-59, Jan 2018.

MAK, A. B. *et al.* Regulation of CD133 by HDAC6 promotes β-catenin signaling to suppress cancer cell differentiation. **Cell Rep**, 2, n. 4, p. 951-963, Out 2012.

MALLARD, B. W.; TIRALONGO, J. Cancer stem cell marker glycosylation: Nature, function and significance. **Glycoconj J**, 34, n. 4, p. 441-452, Ago 2017.

MATSUI, W. H. Cancer stem cell signaling pathways. **Medicine (Baltimore)**, 95, n. 1 Suppl 1, p. S8-S19, Set 2016.

MCGRANAHAN, N.; SWANTON, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. **Cell**, 168, n. 4, p. 613-628, Fev 2017.

MELE, L.; LICCARDO, D.; TIRINO, V. Evaluation and Isolation of Cancer Stem Cells Using ALDH Activity Assay. **Methods Mol Biol**, 1692, p. 43-48, Out 2018.

MEREITER, S. *et al.* O-glycan truncation enhances cancer-related functions of CD44 in gastric cancer. **FEBS Lett**, 593, n. 13, p. 1675-1689, Jul 2019.

MI, W. *et al.* O-GlcNAcylation is a novel regulator of lung and colon cancer malignancy. **Biochim Biophys Acta**, 1812, n. 4, p. 514-519, Abr 2011.

MILLER, K. D. *et al.* Cancer treatment and survivorship statistics, 2019. **CA Cancer J Clin**, 69, n. 5, p. 363-385, Set 2019.

MINISTÉRIO DA SAÚDE. **Estimativa 2020: incidência de câncer no Brasil**. Rio de Janeiro: Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA). Coordenação de Prevenção e Vigilância, 2019

MOKARRAM, P. *et al.* New frontiers in the treatment of colorectal cancer: Autophagy and the unfolded protein response as promising targets. **Autophagy**, 13, n. 5, p. 781-819, Mai 2017.

MOTARI, E. *et al.* Analysis of Recombinant CD24 Glycans by MALDI-TOF-MS Reveals Prevalence of Sialyl-T Antigen. **Am J Biomed Sci**, 1, n. 1, p. 1-11, Jan 2009.

MUNKLEY, J.; ELLIOTT, D. J. Hallmarks of glycosylation in cancer. **Oncotarget**, 7, n. 23, p. 35478-35489, Jun 2016.

MURATA, K. *et al.* Expression of N-acetylglucosaminyltransferase V in colorectal cancer correlates with metastasis and poor prognosis. **Clin Cancer Res**, 6, n. 5, p. 1772-1777, Mai 2000.

NAXEROVA, K. *et al.* Origins of lymphatic and distant metastases in human colorectal cancer. **Science**, 357, n. 6346, p. 55-60, Jul 2017.

NETWORK, C. G. A. Comprehensive molecular characterization of human colon and rectal cancer. **Nature**, 487, n. 7407, p. 330-337, Jul 2012.

NIE, H.; YI, W. O-GlcNAcylation, a sweet link to the pathology of diseases. **J Zhejiang Univ Sci B**, 20, n. 5, p. 437-448, Mai 2019.

ORIAN-ROUSSEAU, V.; SLEEMAN, J. CD44 is a multidomain signaling platform that integrates extracellular matrix cues with growth factor and cytokine signals. **Adv Cancer Res**, 123, p. 231-254, 2014.

PASCHALL, A. V. *et al.* CD133+CD24^{lo} defines a 5-Fluorouracil-resistant colon cancer stem cell-like phenotype. **Oncotarget**, 7, n. 48, p. 78698-78712, Nov 2016.

PASTRANA, E.; SILVA-VARGAS, V.; DOETSCH, F. Eyes wide open: a critical review of sphere-formation as an assay for stem cells. **Cell Stem Cell**, 8, n. 5, p. 486-498, Mai 2011.

PATIL, S. A. *et al.* Overexpression of $\alpha 2, 3$ sialyl T-antigen in breast cancer determined by miniaturized glycosyltransferase assays and confirmed using tissue microarray immunohistochemical analysis. **Glycoconjugate journal**, 31, n. 6-7, p. 509-521, 2014.

PETRETTI, T. *et al.* Altered mRNA expression of glycosyltransferases in human colorectal carcinomas and liver metastases. **Gut**, 46, n. 3, p. 359-366, Mar 2000.

PHUEAOUAN, T. *et al.* Aberrant O-GlcNAc-modified proteins expressed in primary colorectal cancer. **Oncol Rep**, 30, n. 6, p. 2929-2936, Dez 2013.

PINHO, S. S.; REIS, C. A. Glycosylation in cancer: mechanisms and clinical implications. **Nat Rev Cancer**, 15, n. 9, p. 540-555, Set 2015.

PINO, M. S.; CHUNG, D. C. The chromosomal instability pathway in colon cancer. **Gastroenterology**, 138, n. 6, p. 2059-2072, Jun 2010.

QIAN, K. *et al.* Transcriptional regulation of O-GlcNAc homeostasis is disrupted in pancreatic cancer. **J Biol Chem**, 293, n. 36, p. 13989-14000, Set 2018.

RHEE, Y. Y.; KIM, K. J.; KANG, G. H. CpG Island Methylator Phenotype-High Colorectal Cancers and Their Prognostic Implications and Relationships with the Serrated Neoplasia Pathway. **Gut Liver**, 11, n. 1, p. 38-46, Jan 2017.

SANSREGRET, L.; VANHAESEBROECK, B.; SWANTON, C. Determinants and clinical implications of chromosomal instability in cancer. **Nat Rev Clin Oncol**, 15, n. 3, p. 139-150, Mar 2018.

SCHEPERS, A. G. *et al.* Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. **Science**, 337, n. 6095, p. 730-735, Ago 2012.

SCHMITTGEN, T. D.; LIVAK, K. J. Analyzing real-time PCR data by the comparative C(T) method. **Nat Protoc**, 3, n. 6, p. 1101-1108, Jun 2008.

SCHNEIDER, F. *et al.* Overexpression of sialyltransferase CMP-sialic acid:Galbeta1,3GalNAc-R alpha6-Sialyltransferase is related to poor patient survival in human colorectal carcinomas. **Cancer Res**, 61, n. 11, p. 4605-4611, Jun 2001.

SCHUBERT, S. A. *et al.* The missing heritability of familial colorectal cancer. **Mutagenesis**, Out 2019.

SCHULTZ, M. J. *et al.* The Tumor-Associated Glycosyltransferase ST6Gal-I Regulates Stem Cell Transcription Factors and Confers a Cancer Stem Cell Phenotype. **Cancer Res**, 76, n. 13, p. 3978-3988, Jul 2016.

SCHWITALLA, S. *et al.* Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. **Cell**, 152, n. 1-2, p. 25-38, Jan 2013.

SELVAKUMARAN, M. *et al.* Autophagy inhibition sensitizes colon cancer cells to antiangiogenic and cytotoxic therapy. **Clin Cancer Res**, 19, n. 11, p. 2995-3007, Jun 2013.

SHIBUE, T.; WEINBERG, R. A. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. **Nat Rev Clin Oncol**, Abr 2017.

SHIMODA, M.; OTA, M.; OKADA, Y. Isolation of Cancer Stem Cells by Side Population Method. **Methods Mol Biol**, 1692, p. 49-59, Out 2018.

SHIMOKAWA, M. *et al.* Visualization and targeting of LGR5+ human colon cancer stem cells. **Nature**, 545, n. 7653, p. 187-192, Mai 2017.

SLAWSON, C.; COPELAND, R. J.; HART, G. W. O-GlcNAc signaling: a metabolic link between diabetes and cancer? **Trends Biochem Sci**, 35, n. 10, p. 547-555, Out 2010.

SOTTORIVA, A. *et al.* A Big Bang model of human colorectal tumor growth. **Nat Genet**, 47, n. 3, p. 209-216, Mar 2015.

STEINBICHLER, T. B. *et al.* Therapy resistance mediated by cancer stem cells. **Semin Cancer Biol**, 53, p. 156-167, Dez 2018.

STIDHAM, R. W.; HIGGINS, P. D. R. Colorectal Cancer in Inflammatory Bowel Disease. **Clin Colon Rectal Surg**, 31, n. 3, p. 168-178, Mai 2018.

SUN, X.; JU, T.; CUMMINGS, R. D. Differential expression of Cosmc, T-synthase and mucins in Tn-positive colorectal cancers. **BMC Cancer**, 18, n. 1, p. 827, Ago 2018.

SUZUKI, Y. *et al.* Multiregion ultra-deep sequencing reveals early intermixing and variable levels of intratumoral heterogeneity in colorectal cancer. **Mol Oncol**, 11, n. 2, p. 124-139, Fev 2017.

SWINDALL, A. F. *et al.* ST6Gal-I protein expression is upregulated in human epithelial tumors and correlates with stem cell markers in normal tissues and colon cancer cell lines. **Cancer Res**, 73, n. 7, p. 2368-2378, Abr 2013.

SZYMURA, S. J. *et al.* NF- κ B upregulates glutamine-fructose-6-phosphate transaminase 2 to promote migration in non-small cell lung cancer. **Cell Commun Signal**, 17, n. 1, p. 24, Mar 2019.

TANG, Z. *et al.* GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. **Nucleic Acids Res**, 45, n. W1, p. W98-W102, Jul 2017.

TERAO, N. *et al.* Fucosylation is a common glycosylation type in pancreatic cancer stem cell-like phenotypes. **World J Gastroenterol**, 21, n. 13, p. 3876-3887, Abr 2015.

THANIKACHALAM, K.; KHAN, G. Colorectal Cancer and Nutrition. **Nutrients**, 11, n. 1, Jan 2019.

THOMAS, D. *et al.* Truncated O-glycans promote epithelial-to-mesenchymal transition and stemness properties of pancreatic cancer cells. **J Cell Mol Med**, 23, n. 10, p. 6885-6896, Out 2019.

TOGAYACHI, A. *et al.* Beta3GnT2 (B3GNT2), a major polylactosamine synthase: analysis of B3GNT2-deficient mice. **Methods Enzymol**, 479, p. 185-204, 2010.

UCHI, R. *et al.* Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. **PLoS Genet**, 12, n. 2, p. e1005778, Fev 2016.

VAJARIA, B. N.; PATEL, P. S. Glycosylation: a hallmark of cancer? **Glycoconj J**, 34, n. 2, p. 147-156, Abr 2017.

VAN DER WAALS, L. M.; BOREL RINKES, I. H. M.; KRANENBURG, O. ALDH1A1 expression is associated with poor differentiation, 'right-sidedness' and poor survival in human colorectal cancer. **PLoS One**, 13, n. 10, p. e0205536, Out 2018.

VARKI, A. *et al.* Essentials of Glycobiology. *In.* 3 ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2015.

VASCONCELOS-DOS-SANTOS, A. *et al.* Hyperglycemia exacerbates colon cancer malignancy through hexosamine biosynthetic pathway. **Oncogenesis**, 6, n. 3, p. e306, Mar 2017.

VASSALLI, G. Aldehyde Dehydrogenases: Not Just Markers, but Functional Regulators of Stem Cells. **Stem Cells Int**, 2019, p. 3904645, Jan 2019.

VERMEULEN, L. *et al.* Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. **Proc Natl Acad Sci U S A**, 105, n. 36, p. 13427-13432, Set 2008.

WANG, J. L. *et al.* CD44v6 overexpression related to metastasis and poor prognosis of colorectal cancer: A meta-analysis. **Oncotarget**, 8, n. 8, p. 12866-12876, Fev 2017.

WEI, Y *et al.* Activation of PI3K/Akt pathway by CD133-p85 interaction promotes tumorigenic capacity of glioma stem cells. **Proc Natl Acad Sci U S A**, 110, n. 17, p. 6829-6834, Abr 2013.

WIELENGA, V. J. *et al.* Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. **Am J Pathol**, 154, n. 2, p. 515-523, Fev 1999.

WORLD HEALTH ORGANIZATION. **Global Health Observatory Geneva: World Health Organization**. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/cancer>>. Acesso em: 11 Nov. 2019

XU, D. *et al.* Increased expression of O-GlcNAc transferase (OGT) is a biomarker for poor prognosis and allows tumorigenesis and invasion in colon cancer. **Int J Clin Exp Pathol**, 12, n. 4, p. 1305-1314, Abr 2019.

YANG, C. *et al.* High expression of GFAT1 predicts poor prognosis in patients with pancreatic cancer. **Sci Rep**, 6, p. 39044, Dez 2016.

YANG, W. *et al.* Expression of Aldehyde Dehydrogenase 1A1 (ALDH1A1) as a Prognostic Biomarker in Colorectal Cancer Using Immunohistochemistry. **Med Sci Monit**, 24, p. 2864-2872, Mai 2018.

YANG, Y. R. *et al.* OGA heterozygosity suppresses intestinal tumorigenesis in Apc(min/+) mice. **Oncogenesis**, 3, p. e109, Jul 2014.

YEHEZKEL, G. *et al.* O-linked β -N-acetylglucosaminylation (O-GlcNAcylation) in primary and metastatic colorectal cancer clones and effect of N-acetyl- β -D-glucosaminidase silencing on cell phenotype and transcriptome. **J Biol Chem**, 287, n. 34, p. 28755-28769, Ago 2012.

ZEILSTRA, J. *et al.* Stem cell CD44v isoforms promote intestinal cancer formation in Apc(min) mice downstream of Wnt signaling. **Oncogene**, 33, n. 5, p. 665-670, Jan 2014.

ZHANG, H. L. *et al.* c-Myc maintains the self-renewal and chemoresistance properties of colon cancer stem cells. **Oncol Lett**, 17, n. 5, p. 4487-4493, Mai 2019.

ZHANG, L.; SHAY, J. W. Multiple Roles of APC and its Therapeutic Implications in Colorectal Cancer. **J Natl Cancer Inst**, 109, n. 8, Ago 2017.

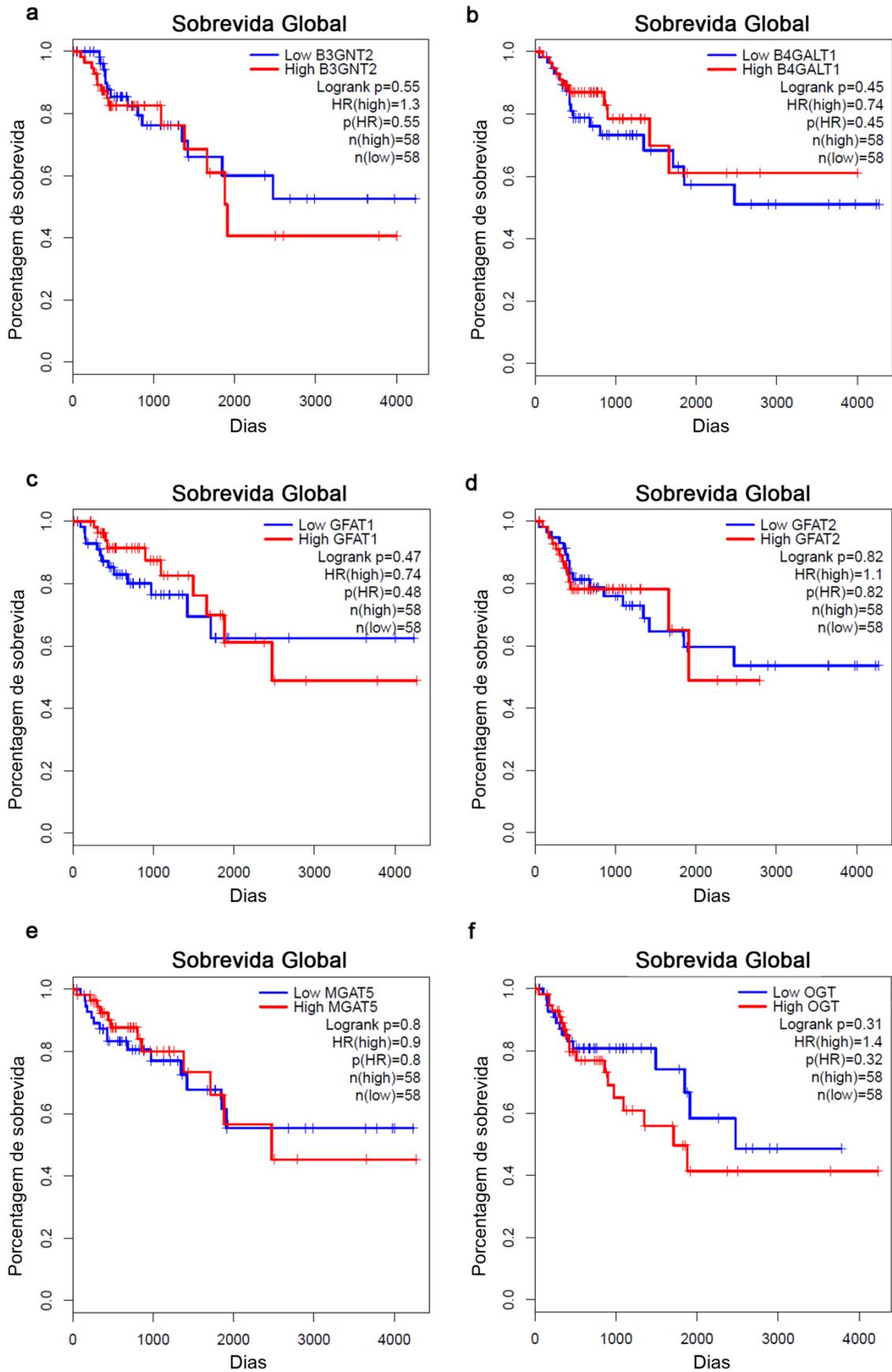
ZHOU, F. *et al.* Alpha2,3-Sialylation regulates the stability of stem cell marker CD133. **J Biochem**, 148, n. 3, p. 273-280, Set 2010.

ZHOU, L. *et al.* Identification of cancer-type specific expression patterns for active aldehyde dehydrogenase (ALDH) isoforms in ALDEFLUOR assay. **Cell Biol Toxicol**, 35, n. 2, p. 161-177, Abr 2019.

ZHUO, J.; WANG, X. Combination of targeting CD24 and inhibiting autophagy suppresses the proliferation and enhances the apoptosis of colorectal cancer cells. **Mol Med Rep**, 20, n. 1, p. 539-548, Jul 2019.

9. Anexos

Anexo 1 – Figuras Complementares de Sobrevida



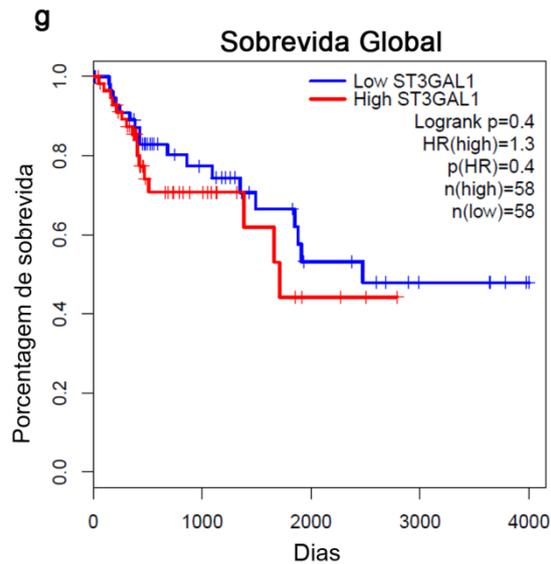
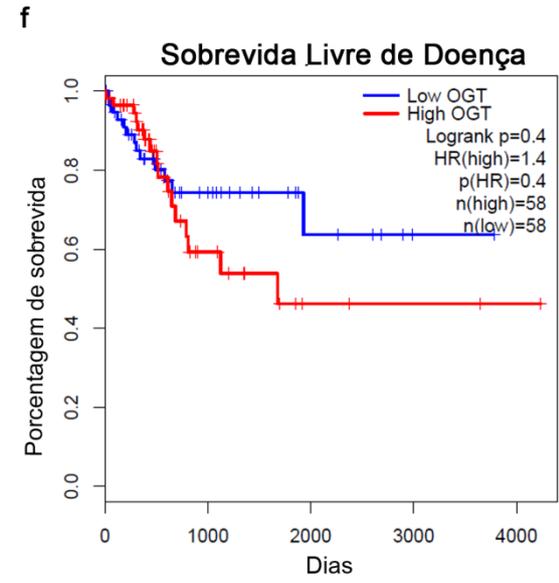
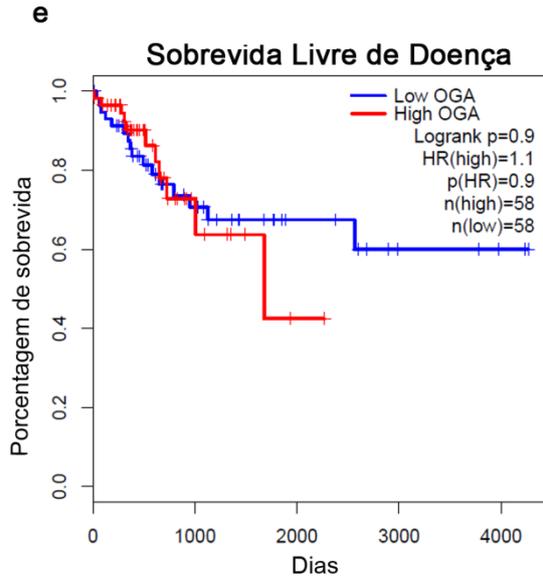
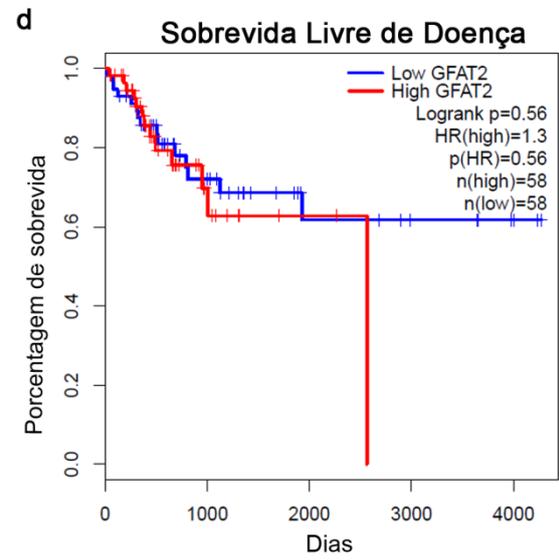
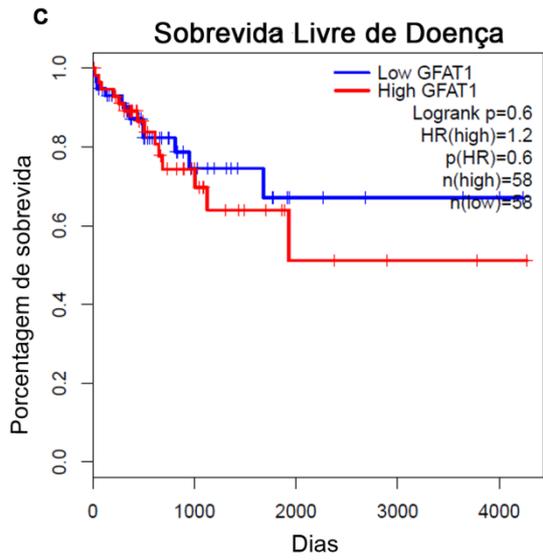
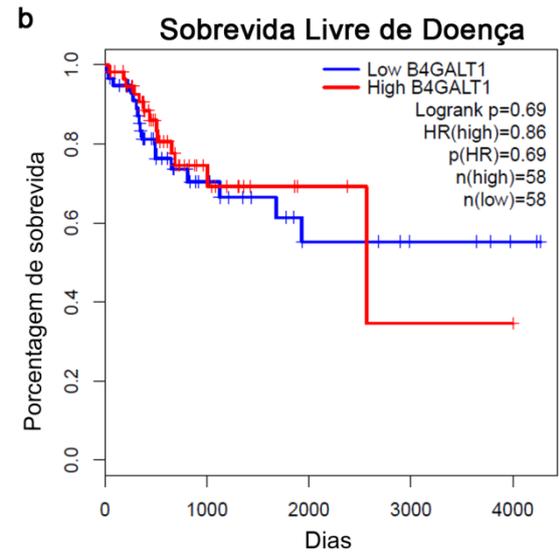
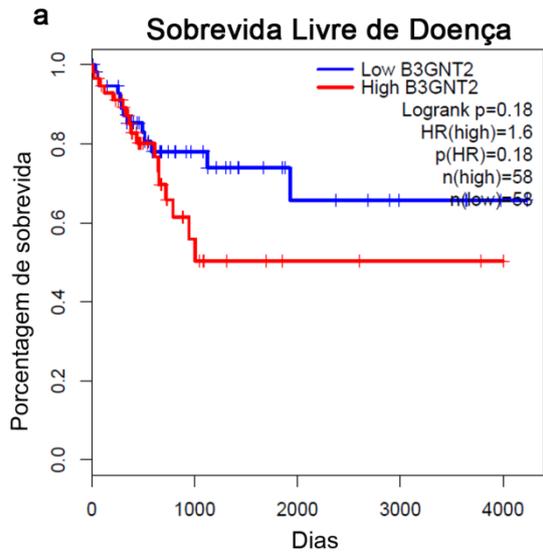


Figura 9.1: Sobrevida global de acordo com os níveis de expressão de glicogenes apenas de pacientes com carcinoma de cólon que não apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados COAD do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSS foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGT* e (g) *ST3GAL1* na sobrevida global de pacientes com câncer de cólon que tem MSS. HR, *Hazard Ratio*; $p(\text{HR})$, valor de p da regressão de Cox usada para cálculo do HR.



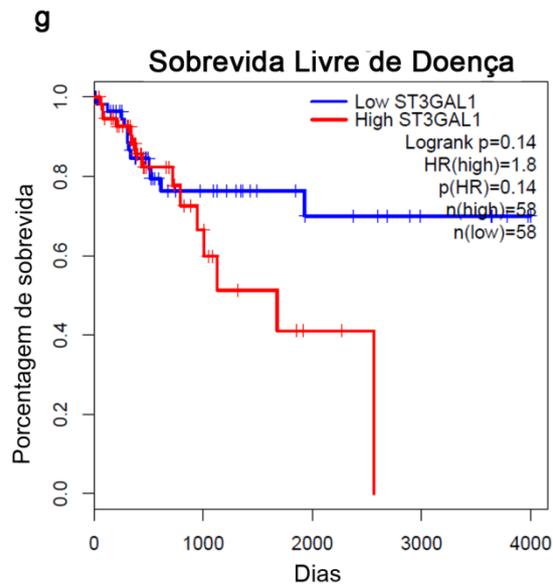
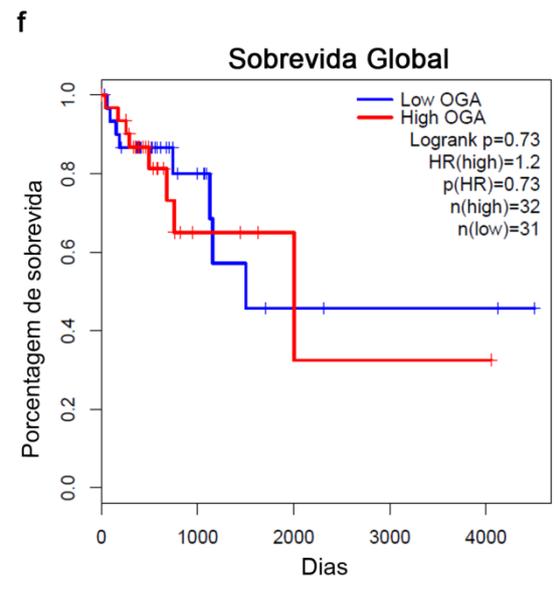
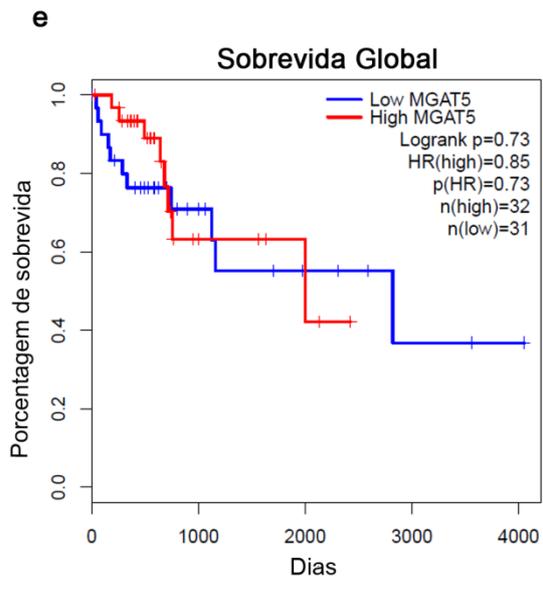
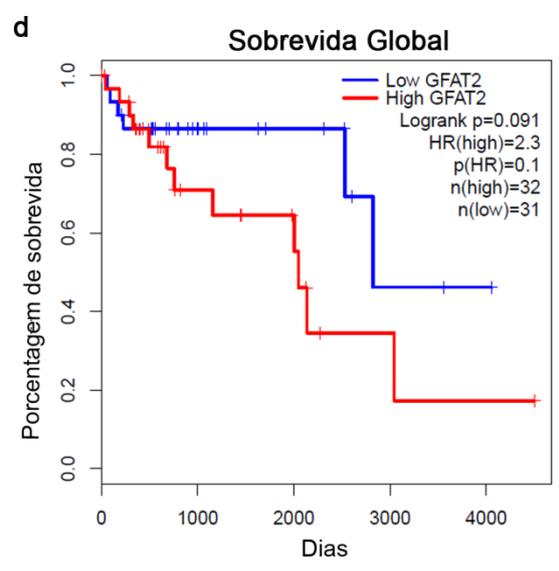
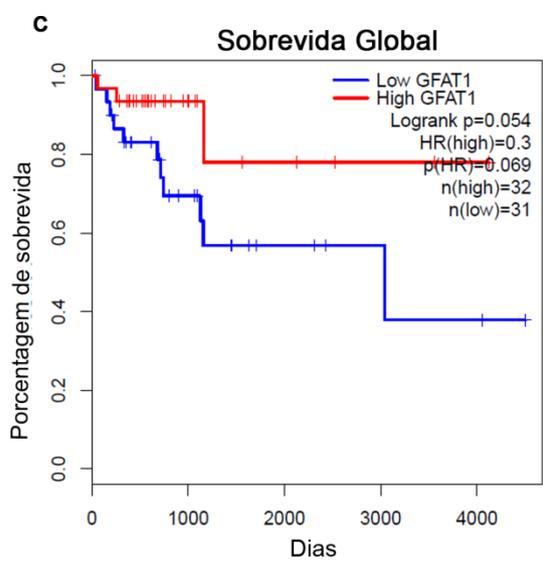
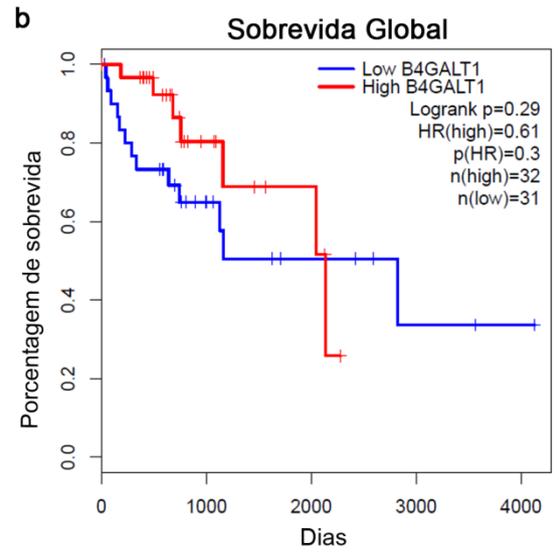
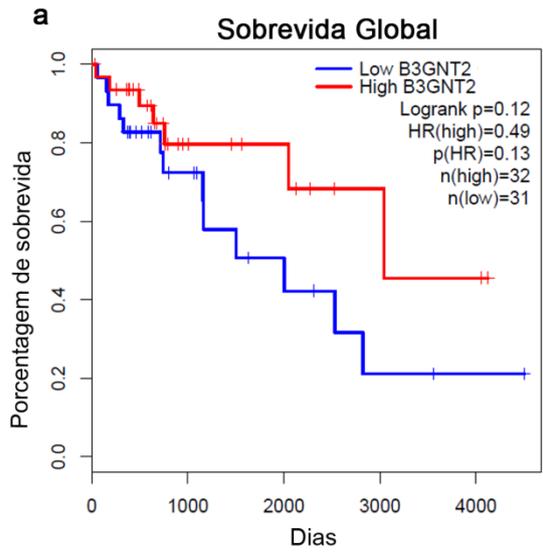


Figura 9.2: Sobrevida livre de doença de acordo com os níveis de expressão de glicogenes apenas de pacientes com carcinoma de cólon que não apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados COAD do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSS foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *OGA*, (f) *OGT* e (g) *ST3GAL1* na sobrevida livre de doença de pacientes com câncer de cólon que tem MSS. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.



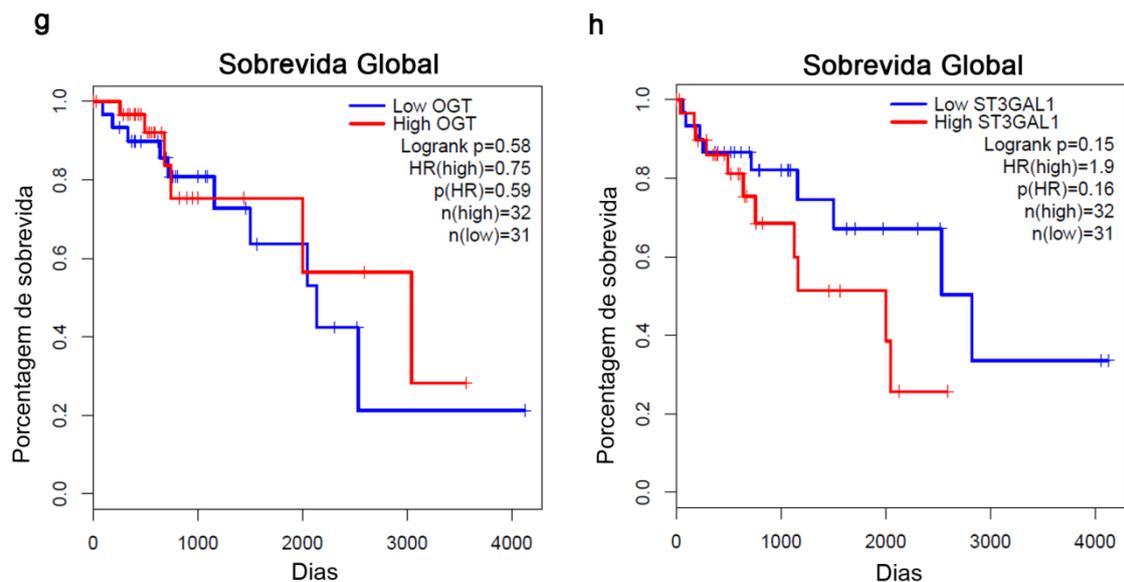
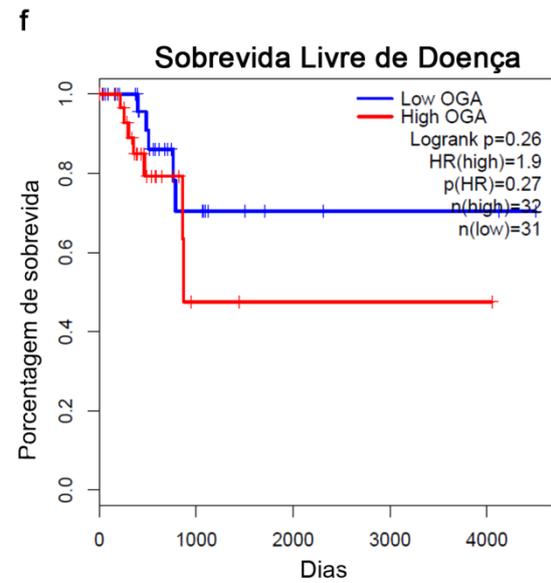
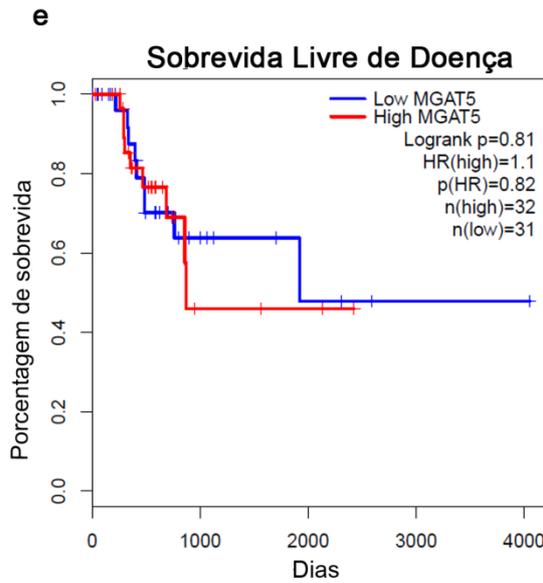
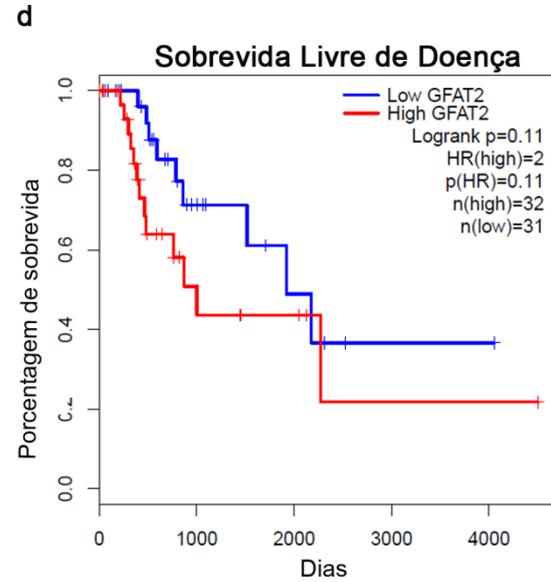
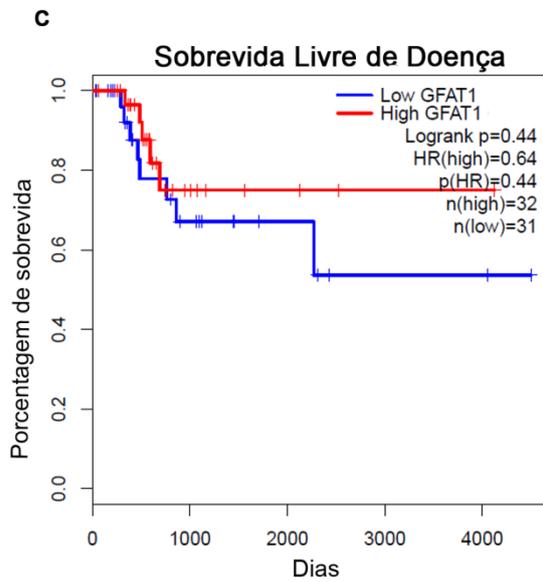
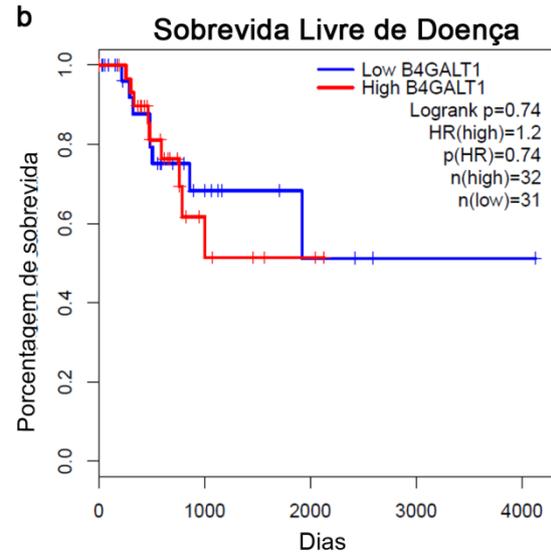
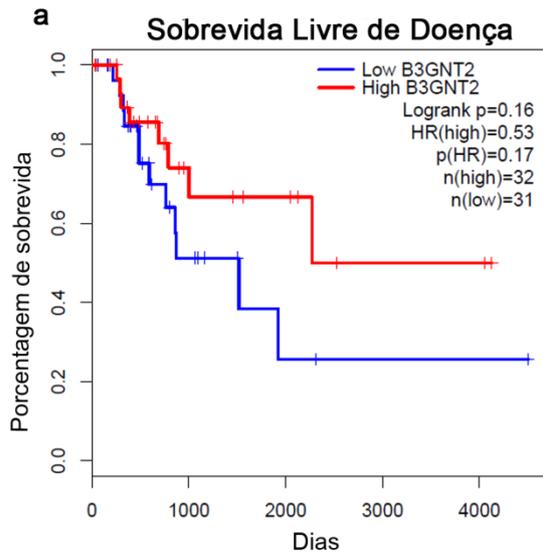


Figura 9.3: Sobrevida global de acordo com os níveis de expressão de glicogenes apenas de pacientes com carcinoma de cólon que apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados COAD do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSI foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevida global de pacientes com câncer de cólon que tem MSI. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.



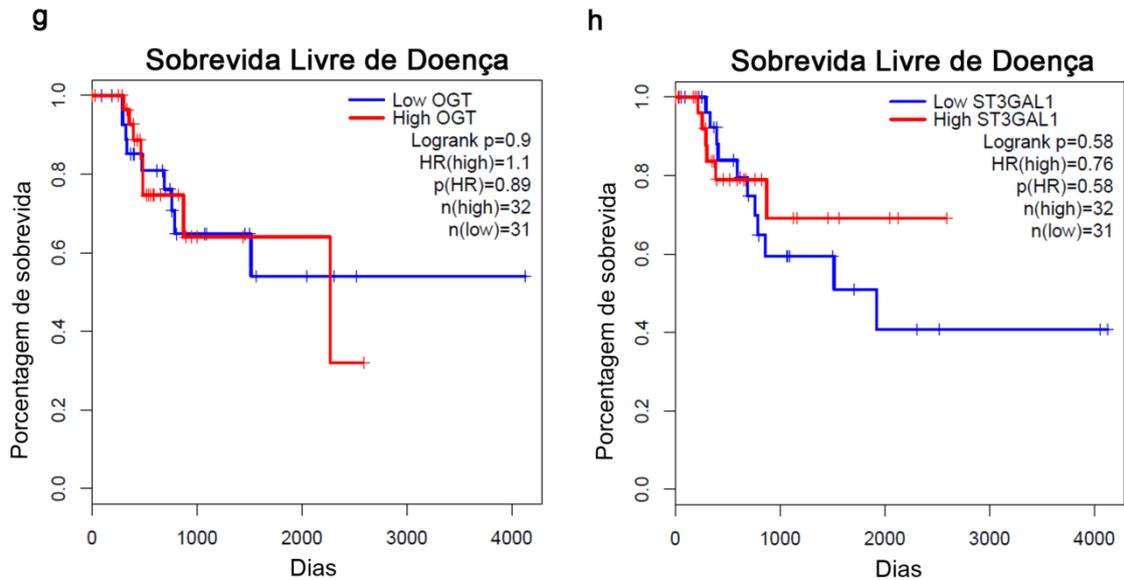
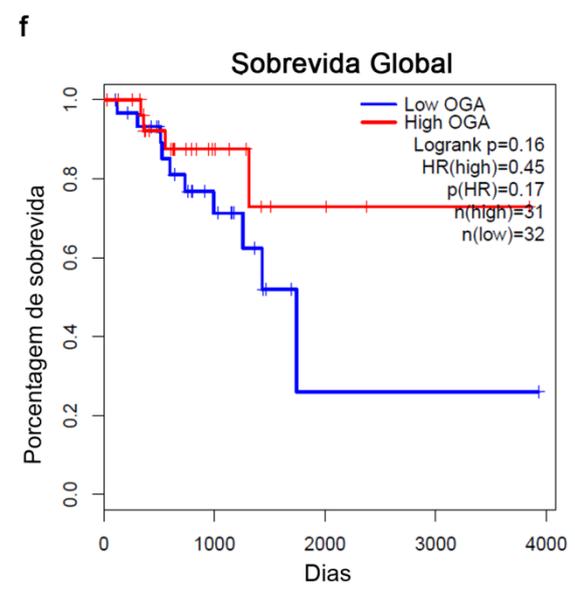
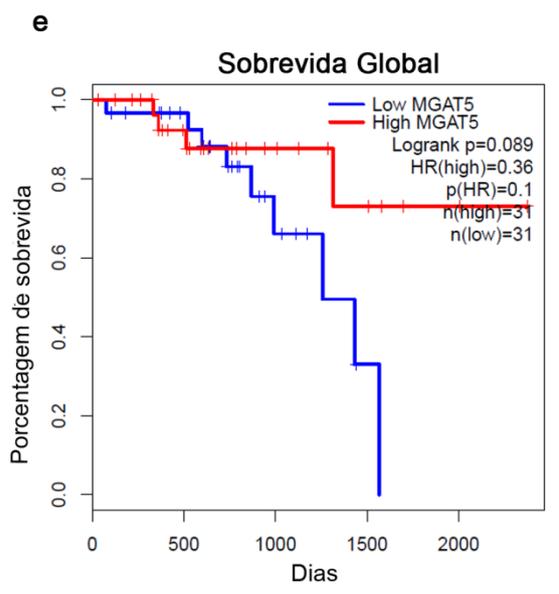
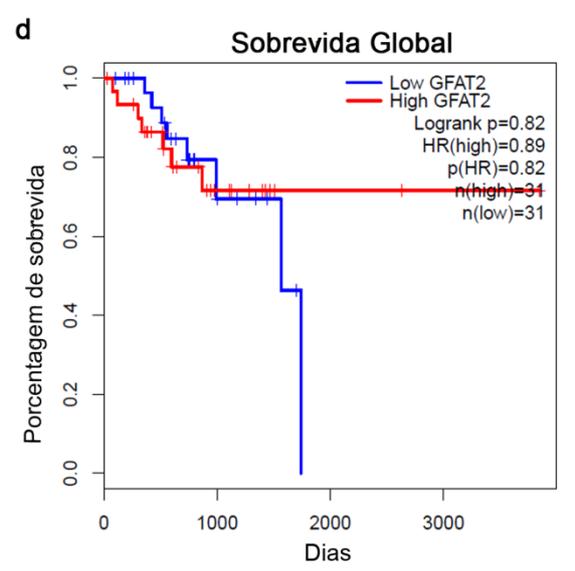
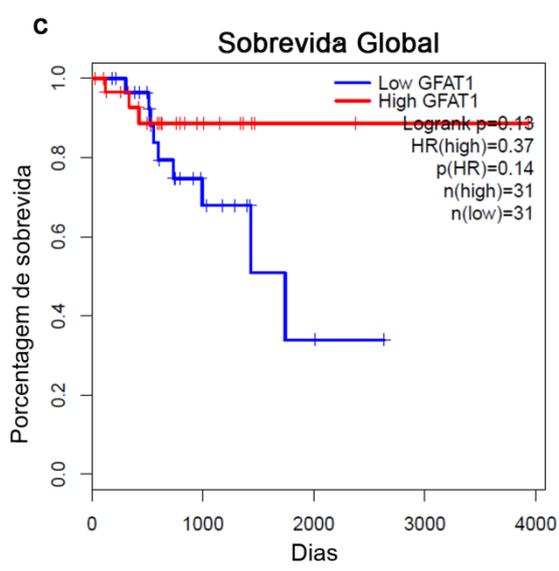
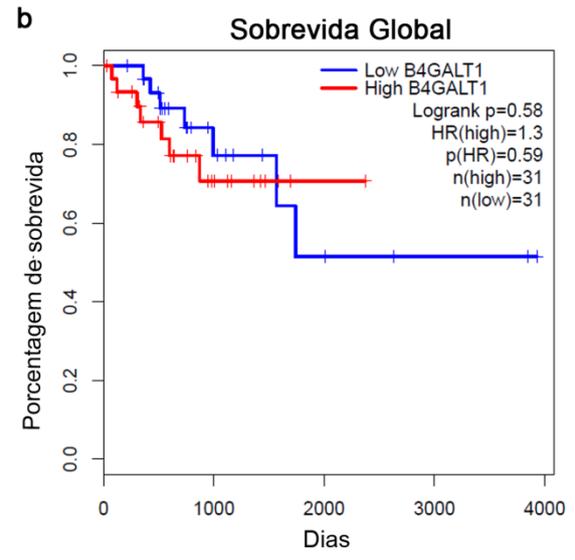
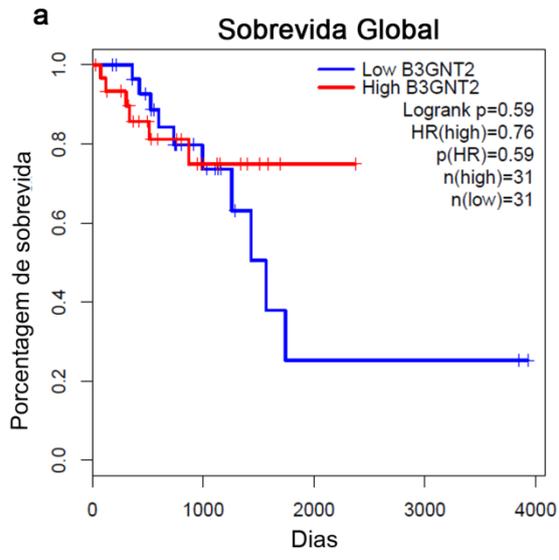


Figura 9.4: Sobrevida livre de doença de acordo com os níveis de expressão de glicogenes apenas de pacientes com carcinoma de cólon que apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados COAD do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSI foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevivida livre de doença de pacientes com câncer de cólon que tem MSI. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.



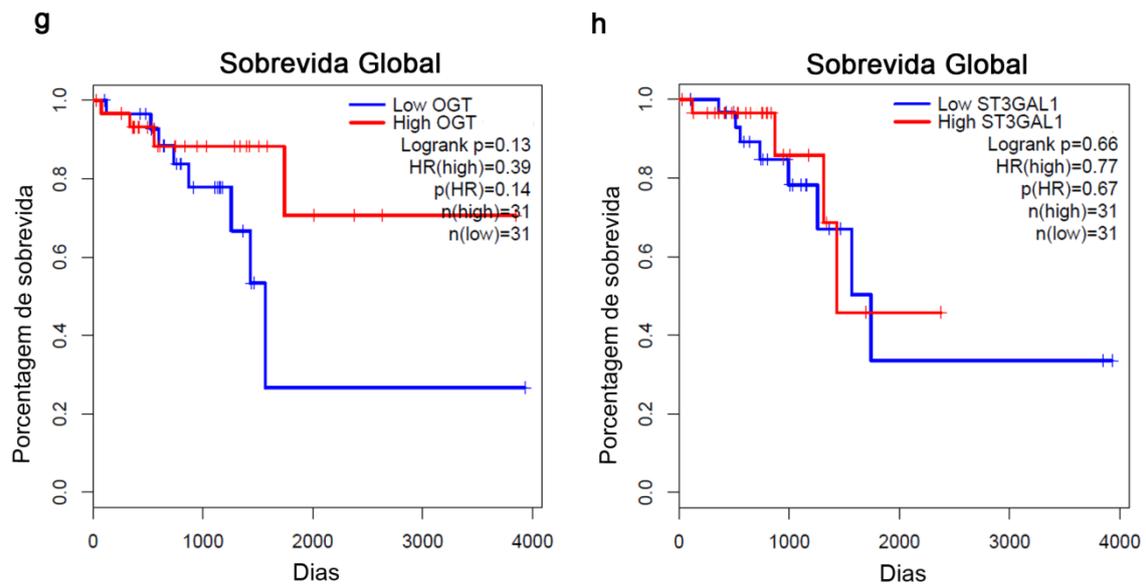
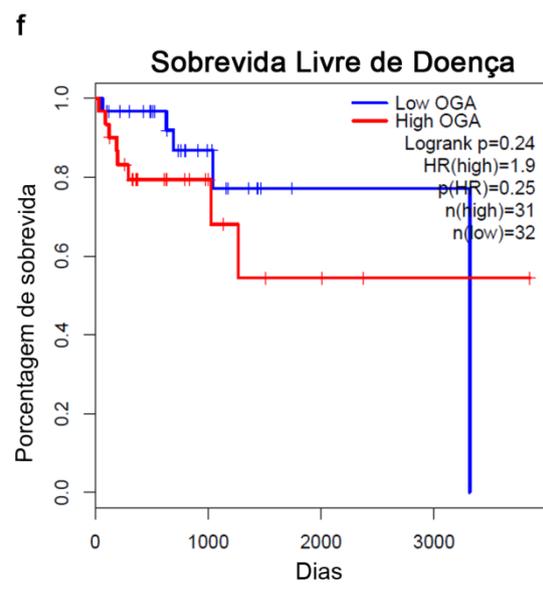
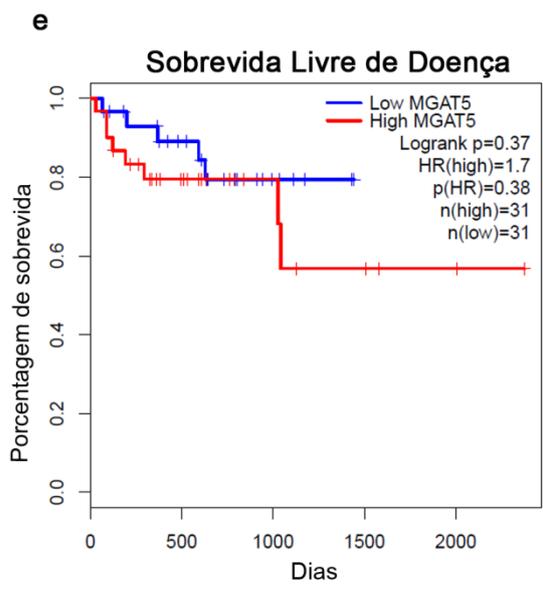
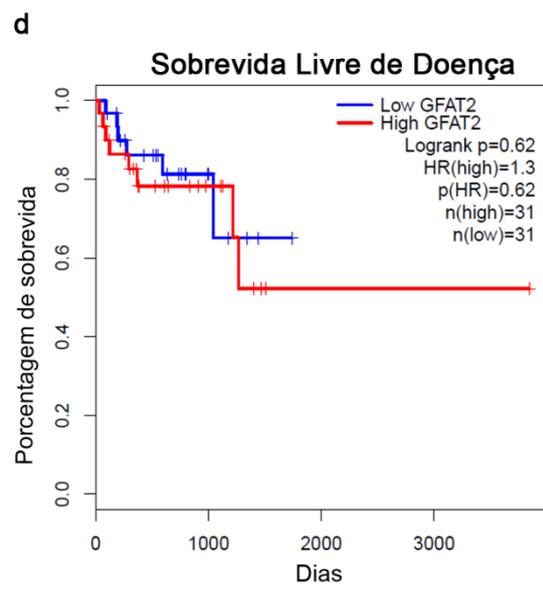
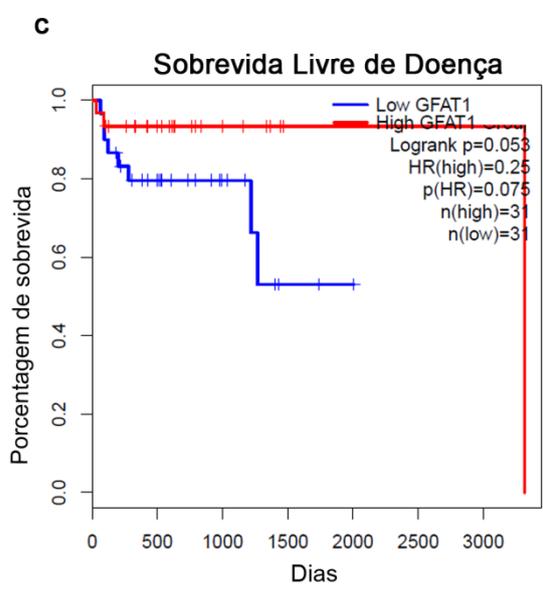
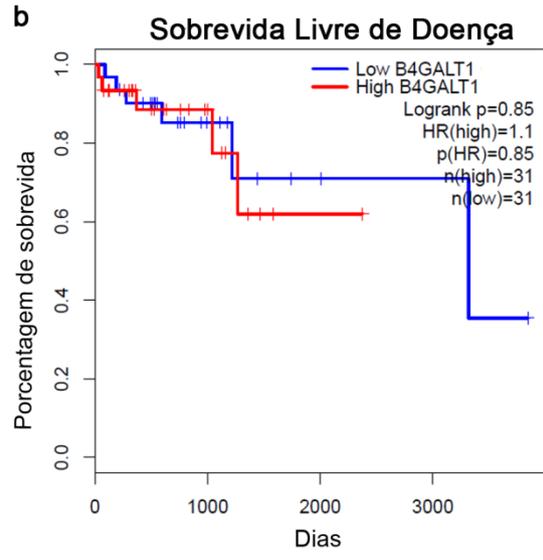
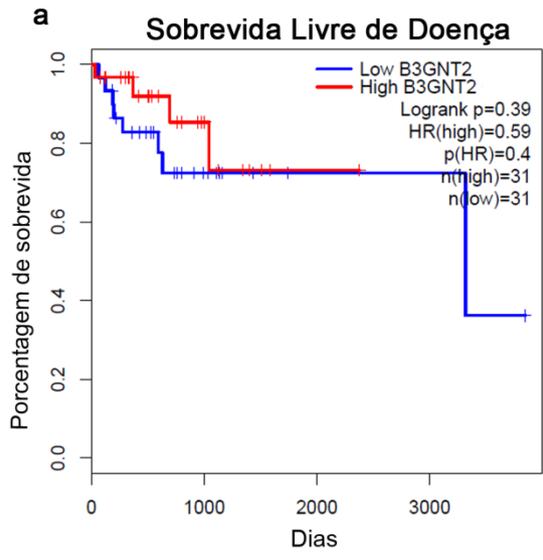


Figura 9.5: Sobrevida global de acordo com os níveis de expressão de diferentes glicogenes em amostras de câncer de reto. Amostras depositadas no banco de dados READ do TCGA foram separadas em tercís de expressão e aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevida global dos pacientes. HR, *Hazard Ratio*; $p(\text{HR})$, valor de p da regressão de Cox usada para cálculo do HR.



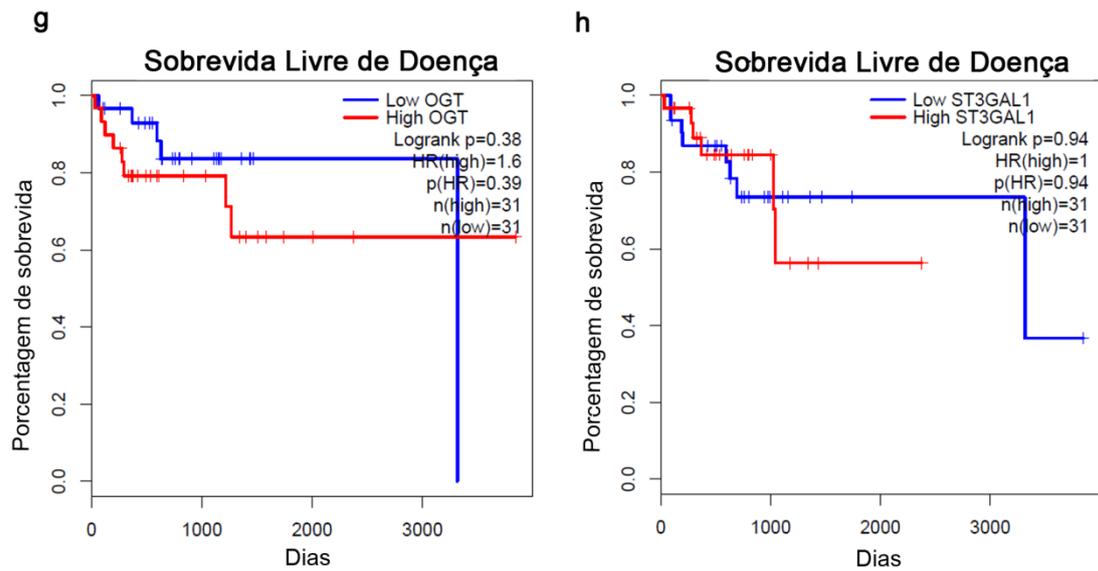
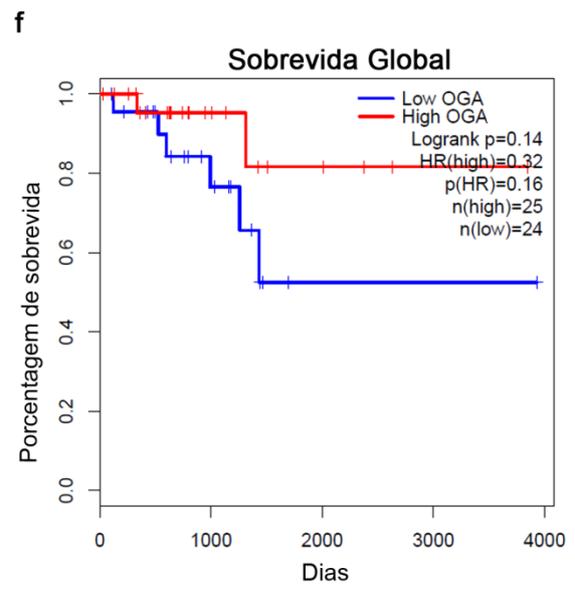
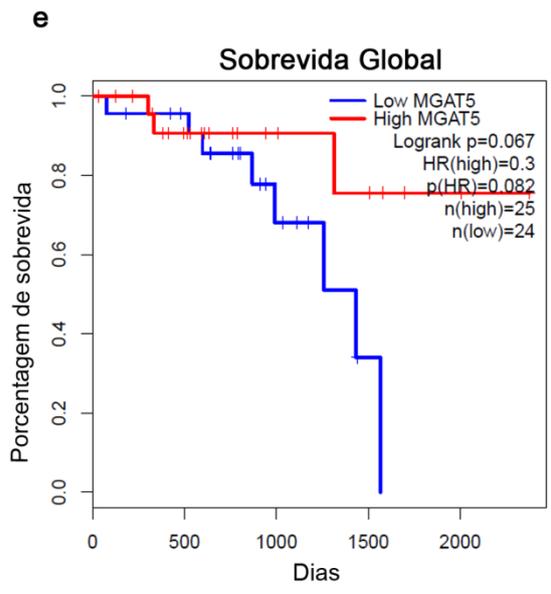
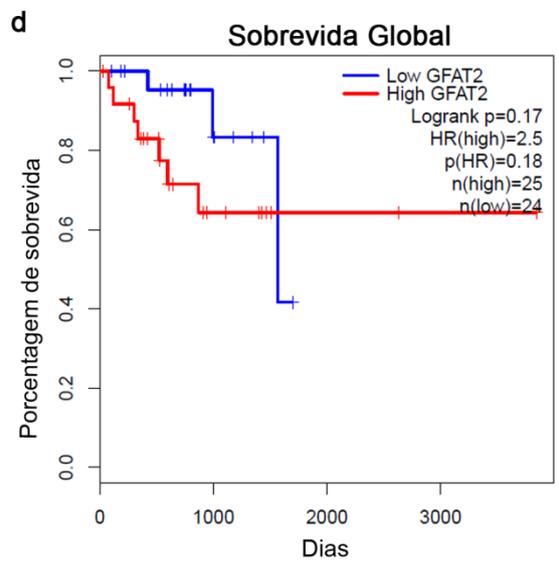
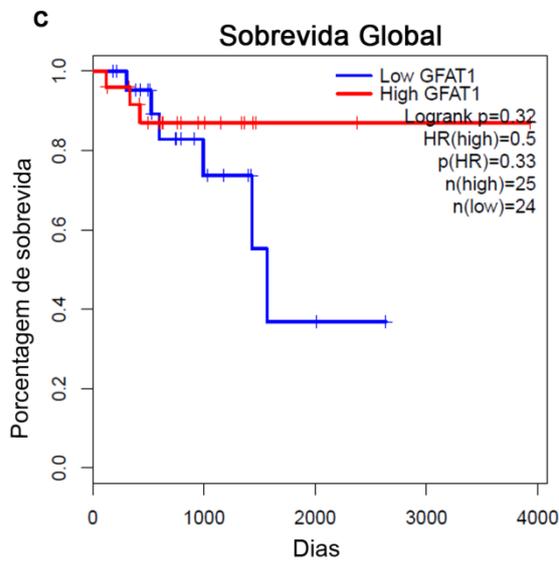
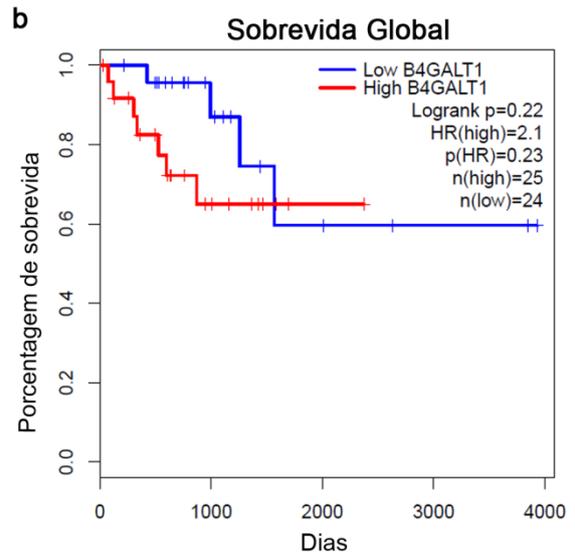
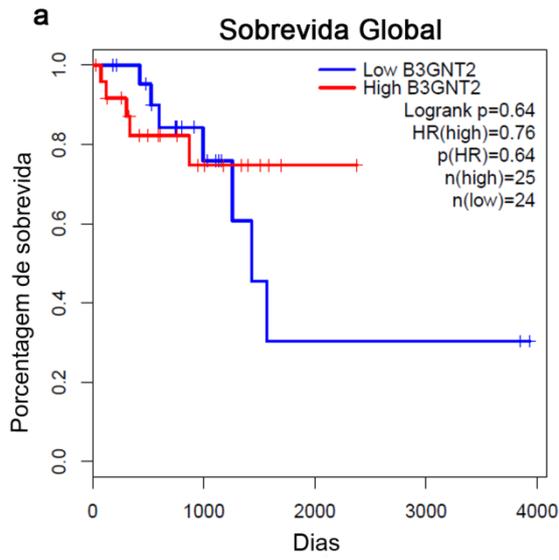


Figura 9.6: Sobrevida livre de doença de acordo com os níveis de expressão de diferentes glicogenes em amostras de câncer de reto. Amostras depositadas no banco de dados READ do TCGA foram separadas em tercís de expressão e aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevivida livre de doença dos pacientes. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.



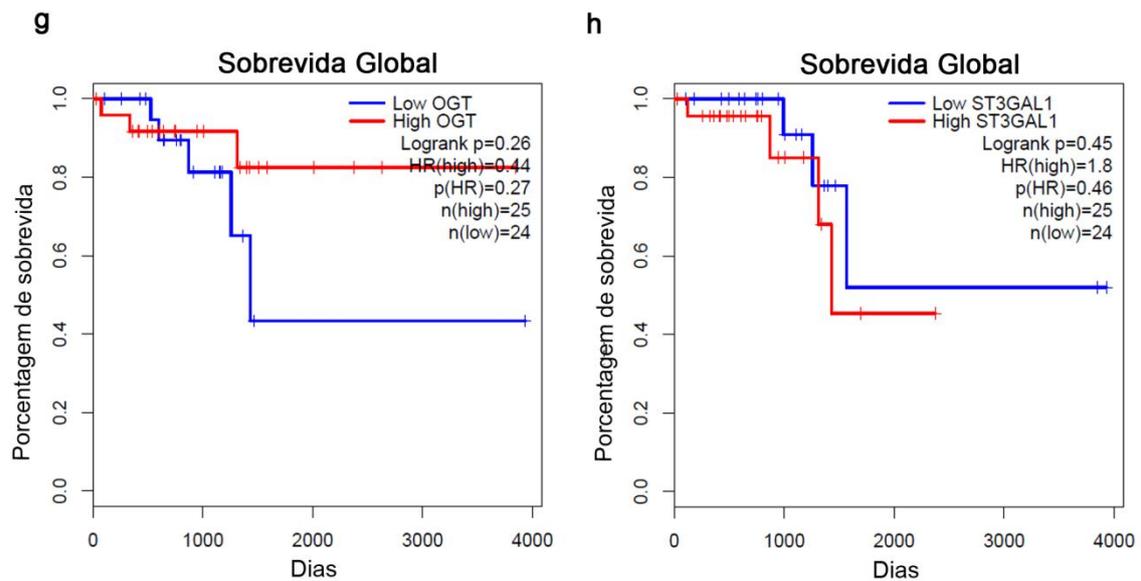
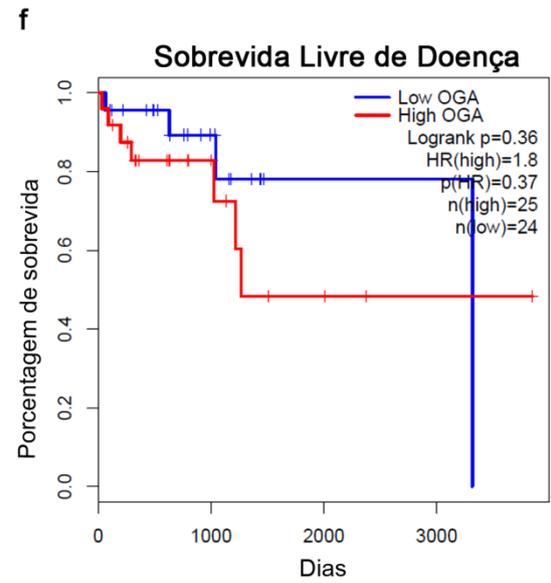
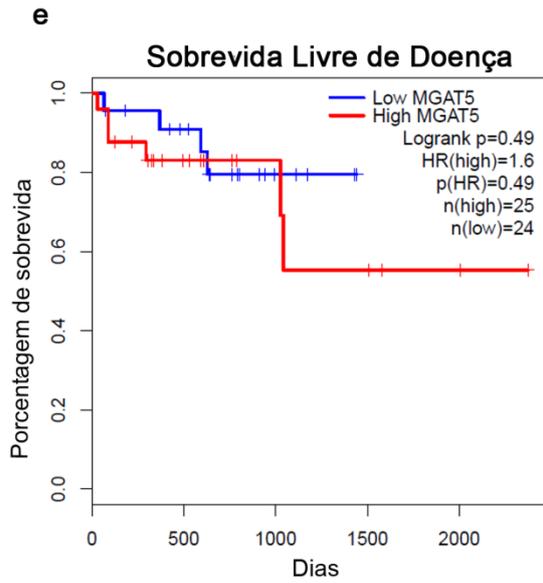
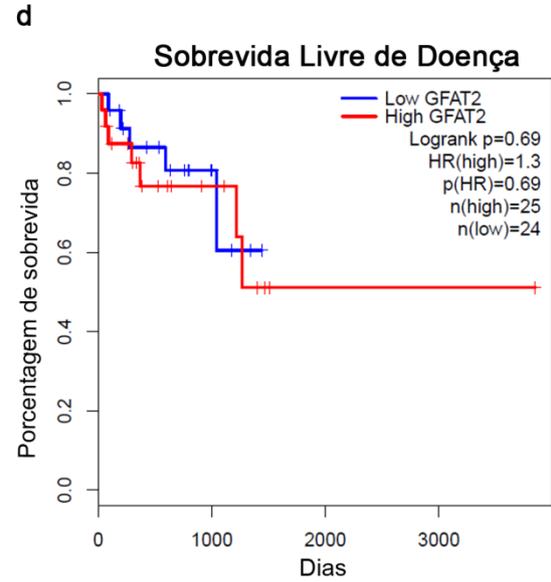
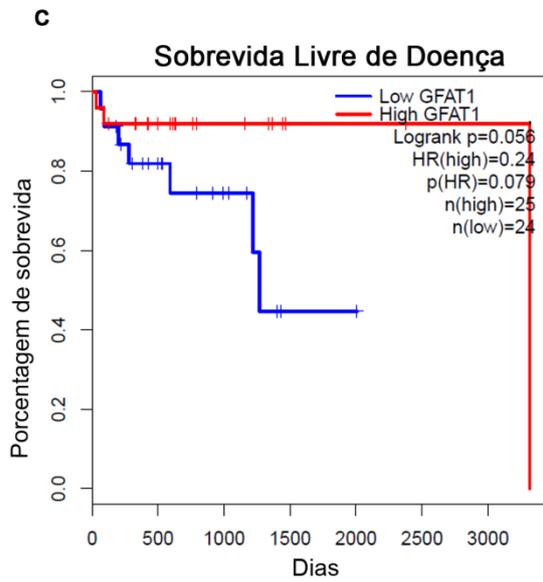
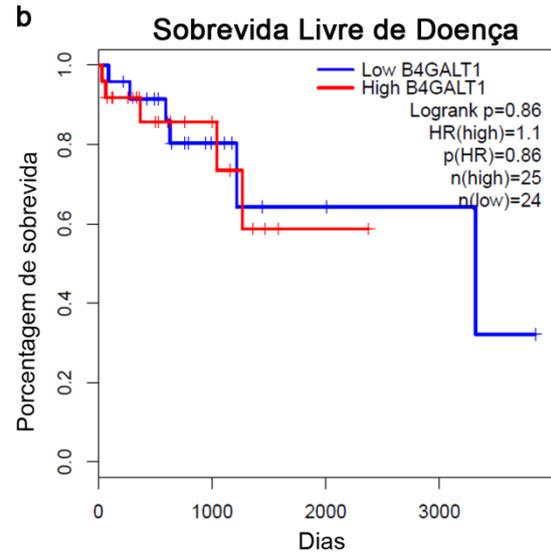
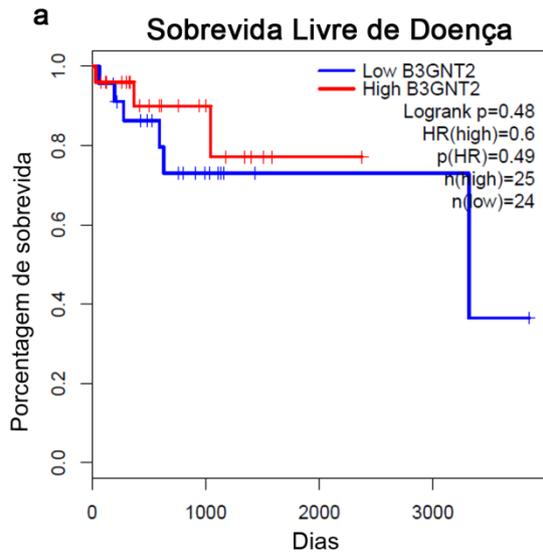


Figura 9.7: Sobrevida global de acordo com os níveis de expressão de glicogenes apenas de pacientes com carcinoma de reto que não apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados READ do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSS foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevida global de pacientes com câncer de reto que tem MSS. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.



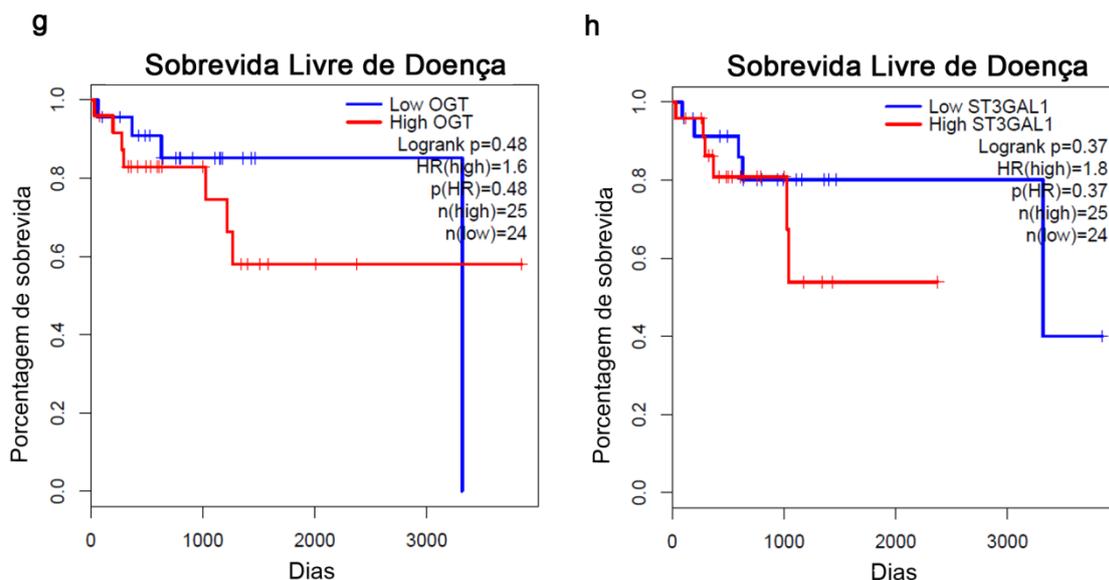
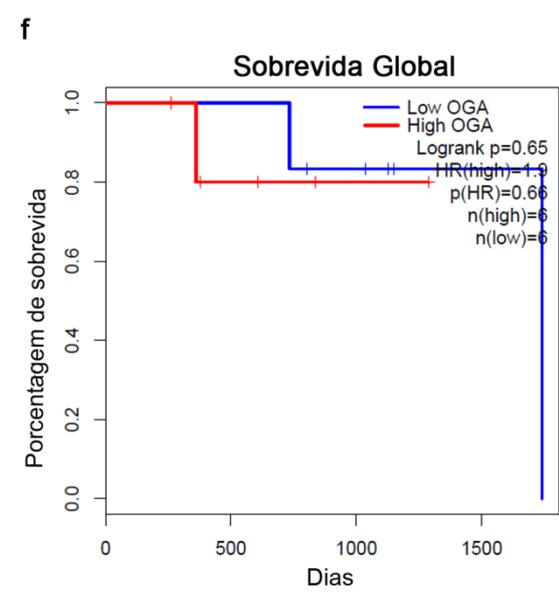
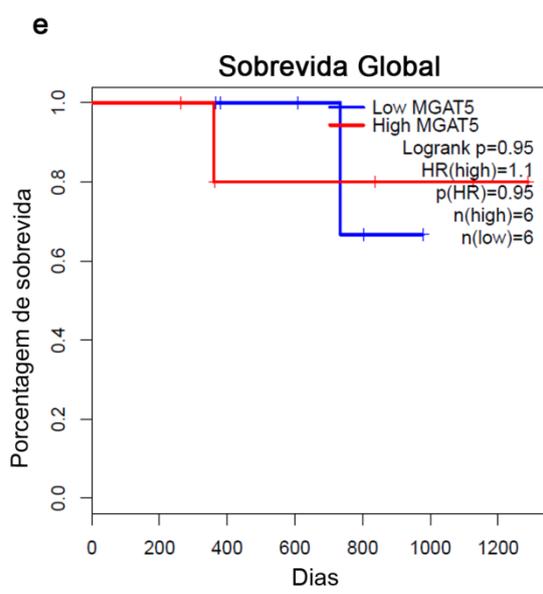
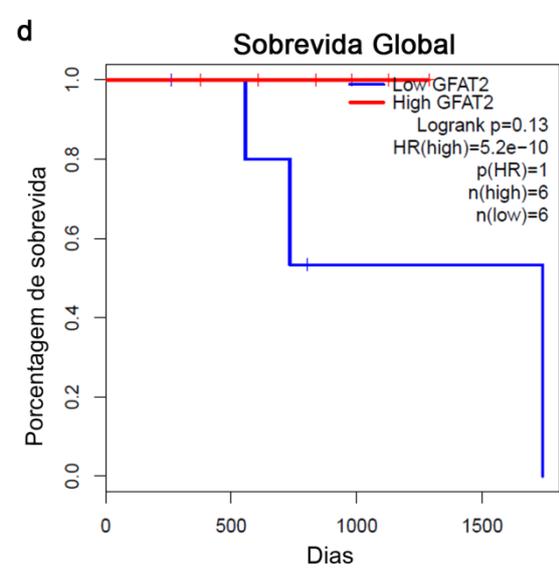
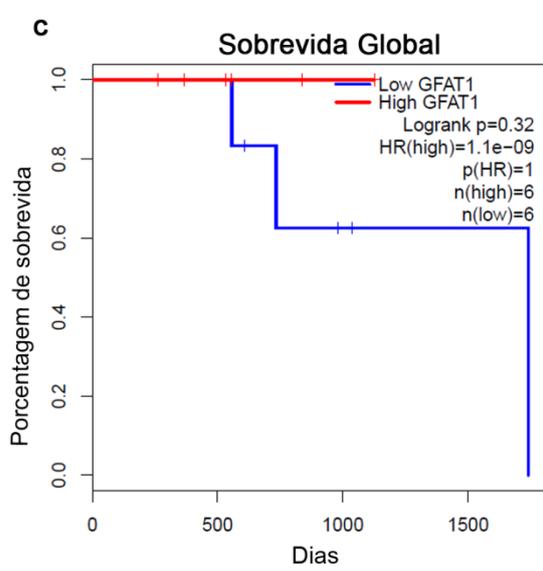
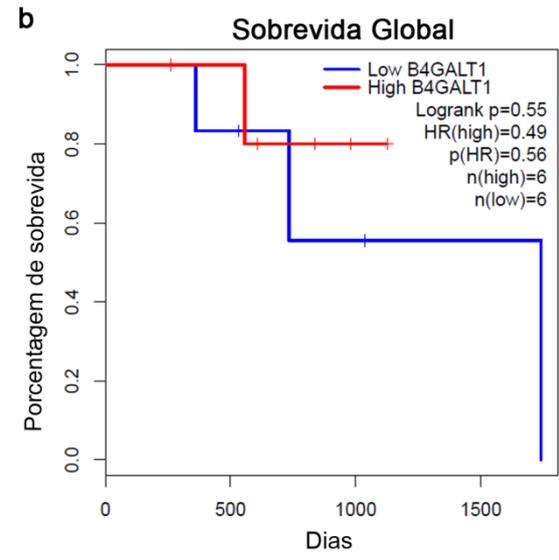
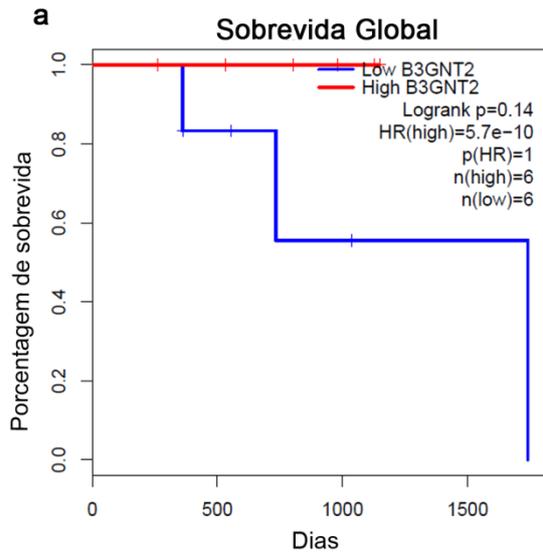


Figura 9.8: Sobrevida livre de doença de acordo com os níveis de expressão de glicogenes apenas de pacientes com carcinoma de reto que não apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados READ do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSS foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevida livre de doença de pacientes com câncer de reto que tem MSS. HR, *Hazard Ratio*; $p(HR)$, valor de p da regressão de Cox usada para cálculo do HR.



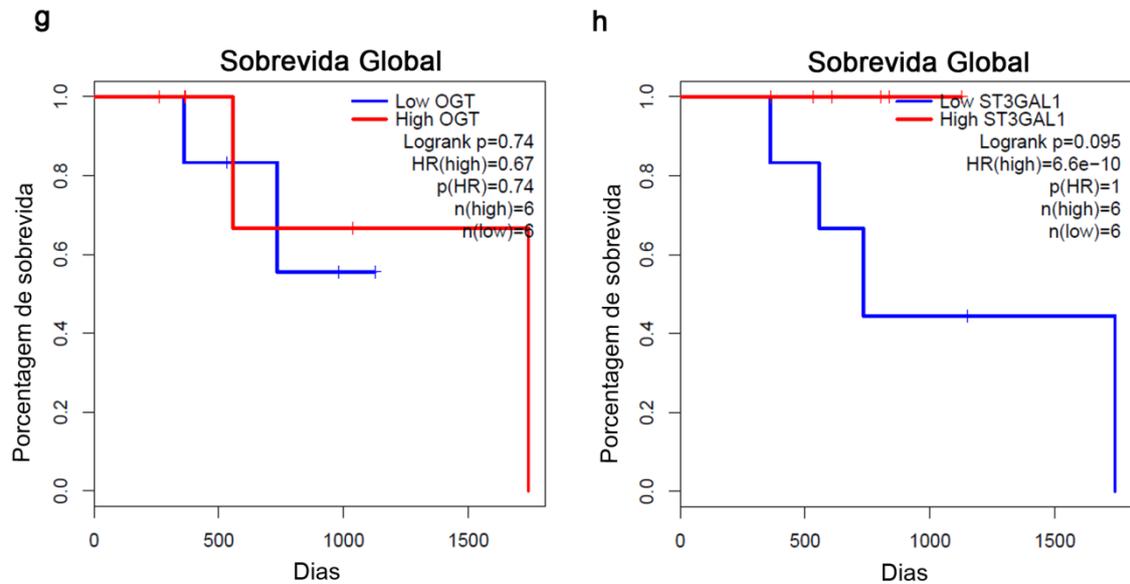
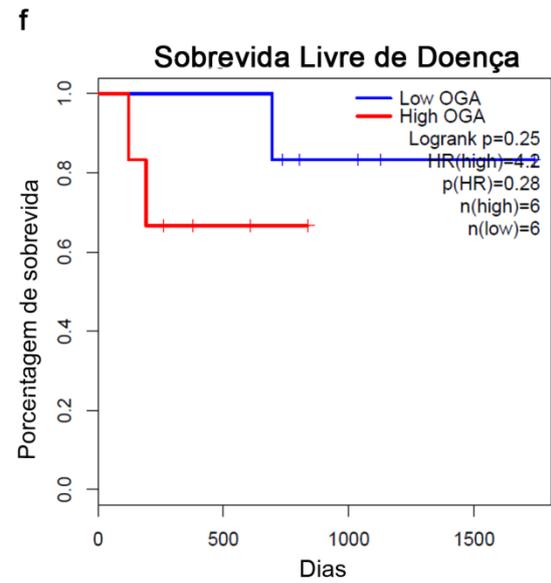
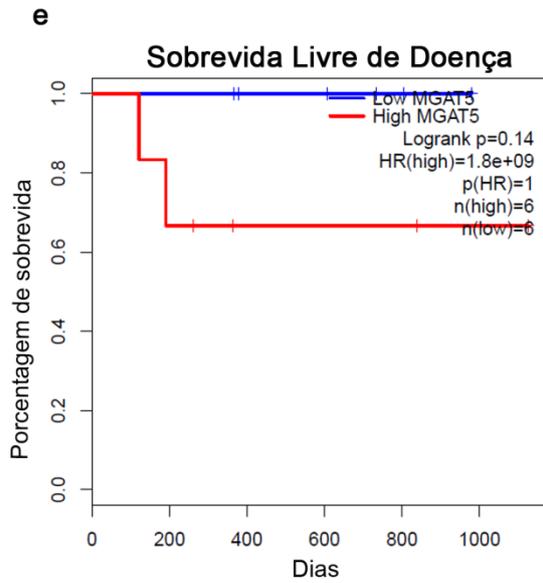
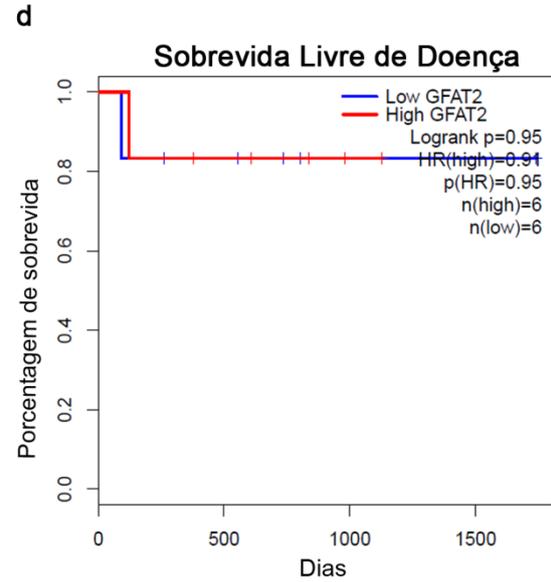
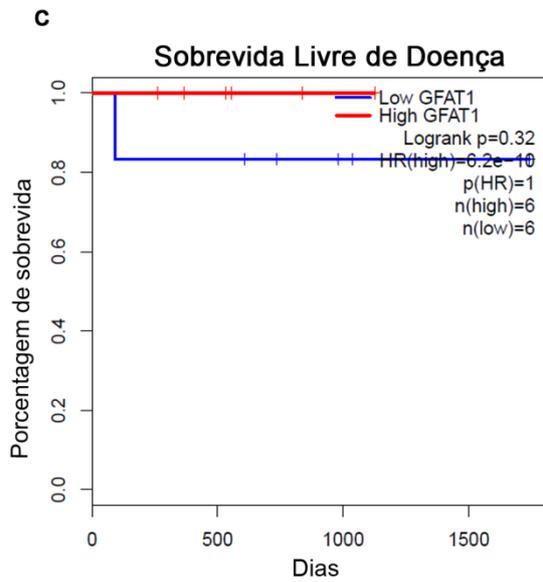
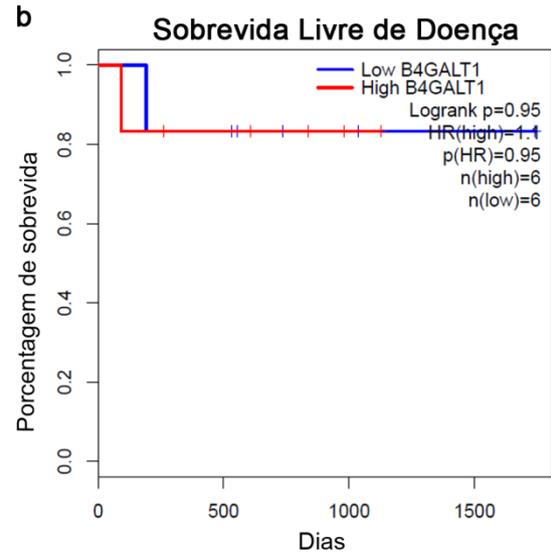
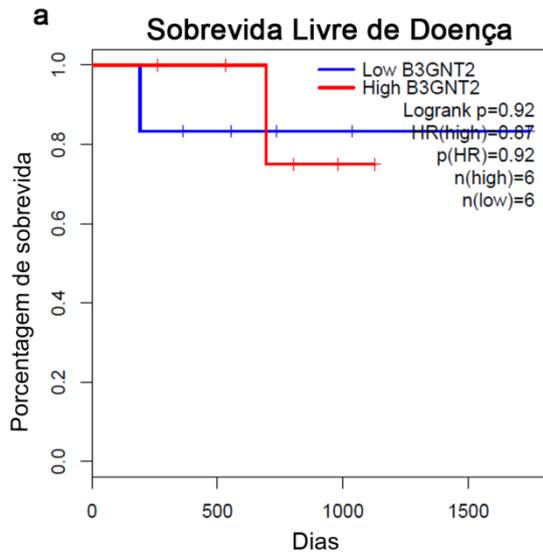


Figura 9.9: Sobrevida global de acordo com os níveis de expressão de glicogenes apenas de pacientes com carcinoma de reto que apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados READ do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSI foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevida global de pacientes com câncer de reto que tem MSI. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.



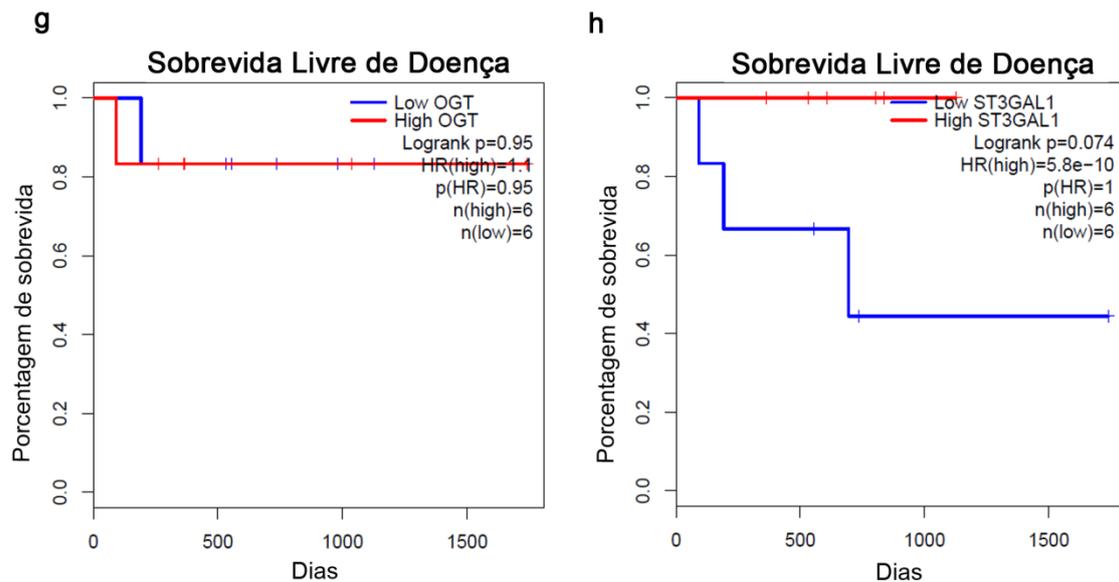


Figura 9.10: Sobrevida livre de doença de acordo com os níveis de expressão de glicogenes apenas de pacientes com carcinoma de reto que apresentam instabilidade de microssatélites. Amostras depositadas no banco de dados READ do TCGA foram separadas de acordo com seu perfil de regiões microssatélite (MSI ou MSS) e apenas as que eram MSI foram separadas de acordo com os níveis de expressão. Aquelas pertencentes ao tercil superior foram consideradas com alta expressão (*high*) enquanto as pertencentes ao tercil inferior foram consideradas com baixa expressão (*low*). Através da plataforma GEPIA foram geradas curvas de Kaplan Meier para avaliar o impacto da expressão de (a) *B3GNT2*, (b) *B4GALT1*, (c) *GFPT1*, (d) *GFPT2*, (e) *MGAT5*, (f) *OGA*, (g) *OGT* e (h) *ST3GAL1* na sobrevida livre de doença de pacientes com câncer de reto que tem MSI. HR, *Hazard Ratio*; p(HR), valor de p da regressão de Cox usada para cálculo do HR.

Anexo 2 e 3 – Artigos científicos submetidos relacionados ao tema Glicobiologia do câncer dos quais participei como co-autora

De: submissions@spandidos-publications.com
Enviado por: submissions=spandidos-publications.com@osats.spandidos-publications.com

Para: jcjunior@inca.gov.br

Data: Quinta-feira, 12 De dezembro De 2019 04:18 PM
Assunto: Spandidos Publications, submission finalized

Dear Dr Julio Freitas,
Thank you for submitting your manuscript.
Your submission has been received successfully, and has been assigned the electronic submission number: 243214
Please click on the link below and log in to access and track the progress of your submission:
<https://www.spandidos-publications.com/osats/>
Your submission will first be checked for completeness and conformance to our guidelines. In the event that any corrections are required prior to evaluation you will be notified shortly. Once your submission has passed our initial checks it will be sent for evaluation by our Editorial Office.
Subsequently, the final decision will be forwarded to you at that time.
Kind regards,

Spandidos Publications

Publication charges for International Journal of Oncology

We offer two publication options: subscription access and Open Access.
Subscription access articles are only available to subscribers for the first 12 months after publication. During this period non-subscribers may purchase individual articles. After 12 months articles become free to view by everyone.

Option 1

Basic charge (for up to 5 pages): €250
Additional pages: €110
Color pages: €390

Option 2

A fixed charge is applied irrespective of the total number of pages or color figures. This option complies with a number of funding bodies.
Open Access: €2,500

Note: this is an automatically generated email. Please do not reply to this email address. Queries can be directed to:
contact@spandidos-publications.com

1 ***N*-glycosylation and RTK signaling affect claudin-3 levels in colorectal cancer cells**

2 Amelia G. Pérez¹, Jéssica Andrade-da-Costa¹, Waldemir F. de Souza¹, Michelle de
3 Souza Ferreira, Mariana Boroni², Ivanir M. de Oliveira³, Carlos A. Freire-Neto¹, Priscila
4 V. Fernandes³, Cristóvão A. de Lanna², Paulo Thiago Souza-Santos⁴, José A.
5 Morgado-Díaz¹, Julio Cesar M. de-Freitas-Junior^{1*}

6 ¹Cellular and Molecular Oncobiology Program, National Cancer Institute (INCA), Rio
7 de Janeiro, Brazil

8 ²Bioinformatics and Computational Biology Laboratory, INCA, Rio de Janeiro, Brazil

9 ³Pathology Division, INCA, Rio de Janeiro, Brazil

10 ⁴Leprosy Laboratory, Oswaldo Cruz Institute, Rio de Janeiro, Brazil

11 *Corresponding author

12 Julio Cesar Madureira de Freitas Junior

13 Address: 37 André Cavalcanti Street, 2th Floor, Rio de Janeiro, RJ, Brazil.

14 Zipcode: 20231-050

15 E-mail: jcjunior@inca.gov.br

16 Phone: +55-21-3207-6599

17 Fax: +55-21-3207-6587

18 **Running title**

19 Pérez et al: *N*-glycosylation and RTK signaling affect claudin-3 levels

20 **Abstract**

21 Colorectal cancer (CRC) is a major public health issue worldwide. Although several
22 studies have been conducted to understand the molecular aspects governing CRC
23 progression, the cellular events leading to tight junction (TJ) destabilization are not
24 completely understood. Considering that claudin-3 is one of the main constituents of TJ,
25 this study aimed to investigate molecular mechanisms involved in the regulation of this
26 protein in CRC cells. First, *in silico* evaluation in each of the CRC consensus molecular
27 subtypes (CMS) revealed that high expression levels of *CLDN3* (gene encoding
28 claudin-3) in CMS2 and CMS3 worsened the patients' long-term survival. Then, we
29 show that *N*-glycan biosynthesis inhibition in CRC cells leads to a decrease in the levels
30 of claudin-3 and, concomitantly, to a reduction in phosphorylation levels of EGFR and
31 IGF1R. We also observed that specific inactivation of these receptor tyrosine kinases
32 (RTKs) also leads to a decrease in claudin-3 levels. We provide evidence that the
33 regulation of claudin-3 levels by changes in RTK signaling can be mediated by PLC and
34 STAT3 in CRC cells. Since RTKs are known to be modulated by *N*-glycans, we
35 evaluated the correlation between *N*-glycogenes and *CLDN3* expression levels in each
36 of the CRC molecular subtypes. CMS1 (MSI immune) was the subtype that
37 concomitantly exhibited the lowest expression of *CLDN3* and *N*-glycogenes (*MGAT5*
38 and *B3GNT8*), whereas CMS2 (canonical) exhibited the highest gene expression levels
39 of *CLDN3* and *N*-glycogenes (*MGAT5*, *ST6GAL1*, and *B3GNT8*). Furthermore, a robust
40 positive correlation between *CLDN3* and *B3GNT8* expression levels was observed in all
41 CMS. These results support the hypothesis of a mechanism integrating RTK signaling
42 and *N*-glycosylation on the regulation of claudin-3 levels in CRC.

43 **Keywords**

44 Receptor tyrosine kinase; claudin-3; *N*-glycans; colorectal cancer; consensus molecular
45 subtypes

46 **Abbreviations**

47 AJC, Apical Junctional Complex; *B3GNT8*, Beta-1,3-*N*-Acetylglucosaminyltransferase
48 8; BRAF, V-Raf Murine Sarcoma Viral Oncogene Homolog B; *CLDN3*, Gene Encoding
49 Claudin-3; CMS, Consensus Molecular Subtypes; CRC, Colorectal Cancer; DPAGT1,
50 Dolichyl-Phosphate *N*-Acetylglucosaminephosphotransferase 1; EGFR, Epidermal

51 Growth Factor Receptor; KRAS, Kirsten Rat Sarcoma Viral Oncogene Homolog;
52 L-PHA, Phytohemagglutinin-L or *Phaseolus vulgaris* Leucoagglutinin; *MGAT5*,
53 Alpha-Mannoside Beta-1,6-N-Acetylglucosaminyltransferase; PLC, Phospholipase C;
54 PP2A, Protein Phosphatase 2; STAT3, Signal Transducer and Activator of Transcription
55 3; *ST6GAL1*, ST6 Beta-Galactoside Alpha-2,6-Sialyltransferase; TJ, Tight Junction.

56 **Introduction**

57 Colorectal cancer (CRC) is one of the most commonly diagnosed cancers in both
58 males and females. Accordingly, this disease is a public health problem, accounting for
59 8.5% of all cancer-related deaths (1). CRC progression occurs through a series of
60 well-defined clinical and histopathological features, ranging from single precursor
61 lesions through benign tumors (serrated or tubular adenoma) to malignant disease (2).

62 CRC is a heterogeneous disease composed of biologically diverse subtypes. One
63 of the most recent gene expression-based subtyping of CRC proposed four consensus
64 molecular subtypes (CMS1-microsatellite instability immune; CMS2-canonical;
65 CMS3-metabolic; and CMS4-mesenchymal), each showing distinguishing molecular
66 disorders related with different clinical outcomes (3, 4). However, consensus molecular
67 classification has not yet been used as a tool to guide clinical decisions, and the constant
68 development of molecular stratification strategies is still necessary to reveal clinical
69 potentials (5).

70 Regardless of the subtype, the disruption of the apical junctional complex (AJC),
71 consisting of tight junctions (TJs) and adherens junctions, is frequently observed during
72 CRC progression (6). The functionality of TJs, which constitute the barrier to the
73 paracellular flow of macromolecules and ions, is regulated among others by the levels
74 of its proteins (7, 8). Changes in subcellular localization and unbalanced (both
75 overexpression and downregulation) claudin levels, which are the main proteins in the
76 regulation of TJs, also affect the stability of TJs in a fine-tuned mechanism (9, 10).
77 Dysregulation of claudin-3 is often observed in CRC (11), and its overexpression
78 destabilizes the TJs, thus increasing the malignancy potential of CRC (12).

79 Previous studies have also demonstrated that the stability of TJs is regulated by
80 glycoproteins, such as RTKs and E-cadherin (13-15), but the regulatory role of *N*-linked
81 glycans in this process is poorly understood. In a pioneering work on the role played by
82 *N*-glycans in TJ stability, it was shown that inhibition of *N*-glycosylation by siRNA to
83 *DPAGT1* (which encodes the enzyme that initiates the synthesis of the dolichol
84 lipid-linked oligosaccharide precursor for protein *N*-glycosylation) promotes the

85 assembly of TJs through the recruitment of the PP2A protein (a negative regulator of TJ
86 biogenesis) to adherens junctions (16).

87 The functionality of RTKs is finely tuned by the *N*-glycans attached on its
88 extracellular domain. Several studies have demonstrated that *N*-glycosylation
89 contributes to ligand binding, kinase activity, and the determination of the receptor's
90 proper conformation (17-20). It has been reported that β 1,6-GlcNAc-branching
91 *N*-glycans (synthesized by MGAT5, also known as GnT-V, see Supplementary material,
92 Fig. S1) on RTKs promote the binding of this branched structure to galectins, thus
93 forming molecular "lattices" that preclude glycoprotein receptor endocytosis, leading to
94 the persistence of cancer-related signaling (21). However, the existence of a mechanism
95 integrating RTK signaling and *N*-glycosylation on the regulation of claudin-3 levels in
96 CRC remains unclear.

97 In the present study, we showed that the high expression levels of *CLDN3* in
98 CRC worsened the patients' long-term survival. Then, we show that *N*-glycan
99 biosynthesis inhibition in CRC cells leads to a decrease in the levels of claudin-3 and,
100 concomitantly, to a reduction in phosphorylation levels of EGFR and IGF1R. We also
101 observed that specific inactivation of these RTKs also leads to a decrease in claudin-3
102 levels. We provide evidence that the regulation of claudin-3 levels by changes in RTK
103 signaling can be mediated by PLC and STAT3 in CRC cells. Since RTKs are known to
104 be modulated by *N*-glycans, we evaluated the correlation between *N*-glycogenes and
105 *CLDN3* expression levels in each of the CRC molecular subtypes. CMS1 (MSI
106 immune) was the subtype that concomitantly exhibited the lowest expression of *CLDN3*
107 and *N*-glycogenes (*MGAT5* and *B3GNT8*), whereas CMS2 (canonical) exhibited the
108 highest gene expression levels of *CLDN3* and *N*-glycogenes (*MGAT5*, *ST6GAL1*, and
109 *B3GNT8*). Furthermore, a robust positive correlation between *CLDN3* and *B3GNT8*
110 expression levels was observed in all four CMS. Altogether, our results corroborate the
111 hypothesis that *N*-glycans play a role in the regulation of RTKs, claudin-3 and TJs,
112 providing a better understanding of CRC biology.

113 **Material and methods**

114 *Chemicals and antibodies*

115 Rabbit monoclonal antibodies anti-pEGFR (Tyr 1068), anti-EGFR, anti-p-IR/p-IGF1R
116 (Tyr 1150, 1151/Tyr 1135, 1136), anti-IGF1R, anti-pAKT (Ser 473), anti-AKT, and

117 anti- α -tubulin were purchased from Cell Signaling Technology (Danvers, MA, USA).
118 Rabbit polyclonal antibodies anti-claudin-3, anti-occludin, and anti-ZO1 were obtained
119 from Thermo Fisher Scientific (Waltham, MA, USA). Anti-mouse GAPDH monoclonal
120 antibody and peroxidase-conjugated anti-mouse were obtained from Santa Cruz
121 Biotechnology (Dallas, TX, USA). Fluorescein-conjugated L-PHA (*Phaseolus vulgaris*
122 *lectin*) lectin was purchased from Vector Laboratories (Burlingame, CA, USA). Alexa
123 Fluor 488-conjugated anti-rabbit was obtained from Molecular Probes (Eugene, OR,
124 USA). Tunicamycin A1 homolog (Tun) was purchased from Sigma-Aldrich. H-89 (PKA
125 inhibitor), and PD 153035 (EGFR inhibitor) were obtained from Sigma-Aldrich (San
126 Luis, MO, USA). Forskolin (PKA activator) and Ly294002 (PI3K inhibitor) were
127 purchased from Sigma-Aldrich. OSI906 (IGF1R inhibitor) was purchased from Selleck
128 Chemicals. PD98059 (MEK1 inhibitor) was purchased from Cell Signaling Technology.
129 U73122 (inhibitor of PLC-dependent processes) was obtained from Cayman Chemical
130 Company (Ann Arbor, MI, USA), while STA-21 (STAT3 inhibitor) was obtained from
131 Santa Cruz. Ruthenium red was purchased from Ted Pella Inc (Redding, CA, USA).

132 *Cell culture and treatments*

133 Colon cancer cells Caco-2, HCT-116, and HT-29 were obtained from the American Type
134 Culture Collection (Manassas, VA, USA). Cells were cultured at 37 °C in a humidified
135 atmosphere of 5% CO₂/air in DMEM (Thermo Fisher Scientific) supplemented with
136 10% heat-inactivated fetal bovine serum (Thermo Fisher Scientific), penicillin G (60
137 mg/l), and streptomycin (100 mg/l). For experimental purposes, cells were seeded on
138 culture flasks, plates, glass coverslips, or on transwell polycarbonate filters with a
139 0.4- μ m pore size (Corning; Corning, NY, USA). After reaching 80% confluence, cells
140 were treated with the drugs for 24 h. The drugs used for the different assays were
141 tunicamycin at 0.25, 0.50, 0.75, and 1 μ g/mL; PD153035 at 1, 10, and 20 μ g/mL;
142 OSI906 at 2, 4 e 8 μ g/mL; U73122 at 7, 14, and 21 μ g/mL; PD98059 at 7, 14, and 21
143 μ g/mL; LY294002 at 2, 4, and 6 μ g/mL; STA-21 at 3, 4, 5, 6, and 7.5 μ g/mL; H-89 at 9
144 μ g/mL; and Forskolin at 4 μ g/mL.

145 *Tissue samples*

146 Well or moderately differentiated human colorectal specimens and mucinous
147 adenocarcinomas were obtained from the surgical resection of 14 Brazilian patients (5
148 males and 9 females/64 ± 10 years) after patient consent. In all cases, control specimens
149 were collected from the accompanying normal mucosa, 5–10 cm away from the
150 carcinoma. All samples were evaluated by a board-certified pathologist. For
151 immunoblotting, the cancer tissue and normal epithelium samples were immediately
152 frozen at -80 °C. The study was carried out with the approval of the National Cancer
153 Institute Ethics Committee (n° 84/04). Clinicopathological features are listed in Table SI
154 (supplementary material).

155 *Western blot*

156 Cell cultures and homogenized tissue samples were washed with phosphate-buffered
157 saline (PBS) and then lysed in a solution containing 1% Triton X-100, 1% NP40, a
158 protease inhibitor cocktail (1 tablet/50 ml buffer; Roche; Basel Switzerland), and a
159 phosphatase inhibitor cocktail (Sigma-Aldrich, 1:100 dilution). Total protein was
160 quantified using a BCA protein assay kit (Bio-Rad; Hercules, California, USA). For
161 Western blotting, samples were subjected to SDS–PAGE, and the separated proteins
162 were transferred to a nitrocellulose membrane. The blots were then probed with primary
163 and peroxidase-conjugated secondary antibodies or biotinylated lectins (Vector
164 Laboratories). The proteins were visualized using an ECL chemiluminescence kit (GE
165 Healthcare; Chicago, IL, USA). Immunoreactive bands from lectin blots were then
166 visualized using the Vector stain ABC kit (Vector Laboratories). The protein or
167 carbohydrate levels were quantified by densitometry using LabWorks 4.6 software
168 (Bio-Rad). The measurements were obtained from sub-exposed photographic films after
169 chemiluminescence reaction, and the values were normalized to the amount of
170 housekeeping (GAPDH or tubulin).

171 *Lectin labeling by flow cytometry*

172 Cells were washed, collected from plates, and fixed with paraformaldehyde (4%) for 8
173 min at room temperature (RT); then, the cells were washed, blocked with bovine serum
174 albumin (4%) (Sigma-Aldrich) for 30 min, and centrifuged at 1500 g for 3 min. Next,
175 fluorescein-conjugated L-PHA lectin was added at a concentration of 5 and 2.5 µg/ml,
176 respectively. After incubation for 20 min at RT, the cells were collected by

177 centrifugation and washed three times with PBS. Finally, 10^4 cells were analyzed by
178 flow cytometer (FASCalibur, Becton-Dickinson; Franklin Lakes, NJ, USA). Unstained
179 cells were used as negative controls for lectin recognition. Fluorescence histograms and
180 median fluorescence data were created and analyzed with CellQuest software
181 (Becton–Dickinson).

182 *Immunofluorescence and confocal microscopy*

183 Cells were grown on glass coverslips until confluence. The cell monolayers were
184 washed with PBS and fixed with methanol for 20 min at $-20\text{ }^{\circ}\text{C}$. Next, the cells were
185 rehydrated with PBS/CM (PBS containing 100 mM of CaCl_2 and 100 mM of MgCl_2 ,
186 pH 8.0) and blocked with 0.2% BSA for 60 min. Subsequently, the cells were incubated
187 overnight with anti-claudin-3 (1:40 dilution), washed with PBS, and incubated for 1 h
188 with Alexa Fluor 488-conjugated rabbit antibody (1:500 dilution). Finally, all coverslips
189 were incubated with DAPI, then washed and assembled using n-propyl gallate
190 (Sigma-Aldrich). The images were acquired using a FV10i-O laser confocal microscope
191 (Olympus; Shinjuku, Tokyo, Japan). The quantification of fluorescence intensity in both
192 the cytoplasm and cell membrane was performed using ICY Bioimage Analysis
193 software.

194 *Transmission electron microscopy*

195 Cells were cultured on transwell polycarbonate filters (Corning) and fixed for 60 min on
196 the apical side of the monolayer with a solution containing 2.5% glutaraldehyde, 1%
197 freshly prepared paraformaldehyde, 8% sucrose, 2 mM of CaCl_2 , and 6 mg/ml of
198 ruthenium red in 0.1 M of cacodylate buffer, pH 7.4. After washing with cacodylate
199 buffer containing ruthenium red for 10 min, they were then postfixed with 1% OsO_4
200 and 6 mg/ml ruthenium red in cacodylate buffer for 45 min. Then, cell monolayers were
201 washed with cacodylate buffer, dehydrated with acetone series, and embedded in Epon
202 resin. Ultrathin sections (70 nm) were stained with lead citrate and observed in a Zeiss
203 CEM-900 transmission electron microscope (Carl Zeiss; Oberkochen, Germany).

204 *CMS subtyping and gene expression level analysis*

205 The CRC molecular subtype classifier (4) was downloaded and applied to 644 primary
206 human colorectal cancer samples using RNA-Seq data from the TCGA (The Cancer
207 Genome Atlas) project database. Clinical-pathological and RNA-Seq data were obtained
208 using the R package TCGABiolinks (22). Then, the expression levels of *CLDN3*, and
209 related glycogenes (*MGAT5*, *ST6GAL1*, and *B3GNT8*) were analyzed among the tumor
210 samples previously classified according to their specific CMS as well as among 51
211 normal samples, also from the TCGA database. Differences in expression levels were
212 evaluated using ANOVA and *t* tests. All analyses and plots were performed in the R
213 environment.

214 *Clinical outcome analysis*

215 To determine the influence of the gene expression levels in clinical outcomes, tumor
216 samples were classified into high or low groups according to the gene expression
217 pattern. The data were divided into three sections (tertiles), and the upper and lower
218 third were considered as high or low groups. Then, the overall survival over six years
219 was analyzed within these low and high expression categories for all samples or by
220 molecular subtype. Survival analyses were carried out using the “survival” package for
221 R.

222 *Statistical analysis*

223 All statistical analyses from *in vitro* assays were performed using the GraphPad Prism 5
224 software (GraphPad Software). Differences were considered statistically significant
225 when *P* was < 0.05. Differences in protein levels were evaluated using ANOVA.

226 **Results**

227 *Identification of colorectal cancer intra-CMS subgroups by expression analysis of* 228 *CLDN3*

229 In a previous study, our research group observed that upregulation of claudin-3
230 increased the malignancy potential of CRC cells (12). Here, by analyzing claudin-3
231 levels in CRC samples and adjacent normal tissue, it was possible to identify two
232 distinct subgroups of tumors, those with high levels and those with low levels of this

233 protein (Fig. 1 A,B). Since gene expression profiling analysis provides clinical outcome
234 predictions for cancer patients (23, 24), we performed an *in silico* analysis using 644
235 CRC samples from the TCGA database to assess the expression levels of *CLDN3*. In
236 order to investigate whether the expression levels of *CLDN3* could affect the overall
237 survival of colorectal cancer patients, the expression values were classified as “high” or
238 “low” using the following strategy: data were divided into tertiles, and the values were
239 defined as “high” or “low” only when they were in the tertiles with the highest or lowest
240 values, respectively. Then, overall survival was calculated according to the expression
241 levels of *CLDN3* in unclassified data (Fig. 1 C), as well as in data classified into the
242 four different molecular subtypes of CRC (Fig. 1 D,E,F,G). Thus, we determined that
243 high expression levels of *CLDN3* in CMS2 and CMS3 worst the patients’ long-term
244 survival (Fig. 1 E,F). Together, these data demonstrate that analysis of *CLDN3*
245 expression was useful for clearly separating the CMS2/CMS3 populations into two
246 groups with distinct clinical outcomes. This data also shows that verification of the
247 expression profile of specific genes within the CRC molecular subtypes represent an
248 appealing strategy for identifying intra-CMS subgroups

249 *Inhibition of N-glycan biosynthesis decreases the protein levels of claudin-3 and*
250 *induces its redistribution in CRC cells*

251 First, we evaluated claudin-3 levels in different colorectal cancer cell lines and found
252 that HCT-116 cells had the highest levels of this protein among the cells analyzed (Fig.
253 2A). On the basis of these results, we chose HCT-116 cells for subsequent analyses.
254 Interestingly, we had previously demonstrated that increased complex *N*-glycan levels
255 in these undifferentiated CRC cells is correlated with a weaker cell–cell adhesion
256 phenotype as compared to differentiated cells (25), corroborating the choice of HCT-116
257 cells to investigate the role of *N*-glycans in TJ stability. We verified that inhibition of
258 *N*-glycosylation with low doses of tunicamycin for 24 h decreased, as expected, the
259 levels of complex *N*-glycans on the surface of the treated cells (Fig. 2B), as verified by
260 labeling with L-PHA lectin (Fig. 2C). Then, we analyzed by Western blot the effect of
261 tunicamycin treatment on TJ component levels (claudin-3, ZO-1, and occludin). We
262 observed that treatment with tunicamycin decreased the claudin-3 levels but did not
263 affect the levels of other evaluated protein constituents (Fig. 2D). It is well known that
264 changes in the protein levels of claudins occur concomitantly with subcellular
265 redistribution (12, 13). Thus, the effect of tunicamycin in the subcellular localization of
266 claudin-3 was evaluated by immunofluorescence. Treatment with tunicamycin promoted
267 a reorganization of claudin-3 subcellular localization characterized by an increase of its

268 levels on the cellular membrane (Fig. 2E, upper panel). Additionally, we investigated
269 whether changes in claudin-3 localization, induced by the inhibition of *N*-glycosylation,
270 could affect the functionality of the TJs. Then, the permeability to ruthenium red dye
271 was evaluated by transmission electron microscopy. Tunicamycin did not fully affect the
272 permeability of ruthenium red, which permeated the paracellular region of monolayers.
273 However, although the full inhibition of permeability was not found, tunicamycin did
274 promote tight cell–cell contacts (Fig. 2E, lower panel), which may contribute to a more
275 differentiated phenotype (25) and to a decrease in invasiveness (26). Together, these
276 data suggest that *N*-glycans are important for regulating not only claudin-3 levels but
277 also the stability of TJ.

278 *Inhibition of N-glycan biosynthesis affects RTK phosphorylation in CRC cells*

279 Previous studies have demonstrated that claudins levels are regulated by signaling
280 pathways activated through transmembrane glycoproteins, such as RTKs (12, 27).
281 However, to our knowledge, the role of RTK *N*-glycosylation in the regulation of its
282 downstream signaling and the possible impact of this mechanism in the claudin-3 levels
283 are not yet understood. Therefore, we evaluated the effects of tunicamycin in the
284 phosphorylation levels of two RTKs (EGFR and IGF1R) and found that inhibition of
285 *N*-glycan biosynthesis decreased the phosphorylation levels of both receptors analyzed
286 (Fig. 3A). These results suggest that modulation of RTKs function by *N*-glycans may
287 impact both receptor functionality and related signaling pathways.

288 *Inhibition of RTK signaling affect the levels of claudin-3*

289 Although tunicamycin has affected both RTKs phosphorylation and claudin-3 levels,
290 inhibition of *N*-glycan biosynthesis has a broad effect and does not only compromise
291 RTKs. So, we decided to evaluate the effects of RTK-specific inactivation on claudin-3
292 levels. HCT-116 cells were treated for 24 h with different concentrations of PD153035
293 (EGFR inhibitor) or OSI906 (IGF1R inhibitor). In addition to promoting a decrease of
294 EGFR and IGF1R phosphorylation levels, these treatments also induced a reduction of
295 claudin-3 levels (Fig. 3 B, C). This result demonstrates that the specific inactivation of
296 EGFR or IGF1R signaling decrease claudin-3 levels. In order to investigate the
297 signaling mechanism through which RTKs could regulate claudin-3 levels, different
298 potentially involved signaling pathways were inhibited using specific inhibitors. Fig. 3D
299 illustrates the possibilities for the regulation of claudin-3 levels, whether directly related
300 to RTKs or not. PLC, MAPK, AKT, and STAT3 (RTK-related signaling) were

301 specifically inhibited using U73122, PD98059, LY294002, and STA-21, respectively.
302 PLC and STAT3 inhibition, but not MAPK and AKT inhibition, led to a significant
303 reduction of claudin-3 levels (Fig. 3 E, F, G, H). Moreover, concomitant treatment with
304 U73122 and STA-21 did not potentiate the inhibitory effect (Fig. 3I). Interestingly, PKA
305 (non-RTK-related signaling) inhibition by H-89 also led to a significant reduction of
306 claudin-3 levels (Fig. 3H). Collectively, these results show that a complex regulatory
307 network, regardless of the involvement of pathways related to RTKs, can influence the
308 protein levels of claudin-3. Furthermore, these data suggest that the regulation of
309 claudin-3 levels due to modulation of RTK activity, including by changes in its
310 *N*-glycosylation pattern, may be associated with disturbances in the PLC and STAT3
311 pathways.

312 *CLDN3* and *B3GNT8* expression correlate positively in CRC

313 Previous studies have shown that changes in *N*-glycosylation affect the functionality of
314 RTKs (17-19). Because of this, we decided to evaluate the correlation between the
315 expression of *CLDN3* and transcript levels of *N*-glycan-related genes in CRC. The three
316 *N*-glycogenes (see Supplementary material, Fig. S1) analyzed were: I) *MGAT5*, the gene
317 encoding human *N*-acetylglucosaminyltransferase V (MGAT5 or GnT-V), which is
318 responsible for the synthesis of β 1,6-GlcNAc branching *N*-glycan structures that has
319 been widely associated with malignant phenotype (28); II) *B3GNT8*, the gene encoding
320 human UDP-GlcNAc:betaGal beta-1,3-*N*-acetylglucosaminyltransferase 8, an enzyme
321 that is involved in the biosynthesis of poly-*N*-acetylglucosamine chains on
322 β 1,6-branched *N*-glycan, which increase the reactivity to L-PHA when overexpressed in
323 CRC cells, thus being potentially involved in malignancy (29); and III) *ST6GAL1*, the
324 gene encoding human ST6 beta-galactosamide alpha-2,6-sialyltransferase 1, a
325 sialyltransferase that adds an α 2-6-linked sialic acid to the *N*-glycan, which the
326 upregulation was reported in CRC (30, 31). Our *in silico* approach revealed a robust
327 positive correlation between *CLDN3* and *B3GNT8* expression levels in all four CMS
328 and also in all stages of CRC (Figs. 4 and 5). A weaker positive correlation was also
329 observed between *CLDN3* and *ST6GAL1* expression in stages II, III and IV of CRC
330 (Fig. 5). Surprisingly, a negative correlation between *CLDN3* and *MGAT5* expression
331 levels was observed in stages I and III, as well as in CMS2 (Figs. 4 and 5). Since

332 *N*-glycans can regulate RTK activity, these results support the existence of a regulatory
333 mechanism that could connect RTKs, *CLDN3*, and *N*-glycan-related glycoenes.
334 However, more studies are needed to elucidate this point, for such purpose subsequent
335 analyses are ongoing.

336

337 *CLDN3* and *N*-glycan-related glycoenes show similar expression patterns within
338 molecular subtypes of colorectal cancer

339 Considering the correlation data shown here, we also investigated the expression profile
340 of *CLDN3* and *N*-glycoenes in different CRC stages and CMS. The *in silico* analysis
341 comparing the molecular subtypes to each other showed that CMS1 is the subtype that
342 concomitantly exhibits the lowest expression levels of *CLDN3*, *MGAT5*, and *B3GNT8*
343 when compared to the other subtypes, while CMS2 is the subtype that concomitantly
344 exhibits the highest expression levels of *CLDN3*, *MGAT5*, *ST6GAL1*, and *B3GNT8*
345 when compared to the other subtypes (Fig. 6, upper panel). Interestingly, the expression
346 levels of *CLDN3*, *ST6GAL1*, and *B3GNT8* were found in CMS1 at lower levels when
347 compared to normal tissue. In addition, all other analyzed genes had increased levels in
348 CMS2 when compared with CMS1. Corroborating previous results showing that the
349 upregulation of *ST6GAL1* is frequently observed in CRC samples (32, 33), here we
350 identified that the expression levels of *ST6GAL1* were significantly increased in CMS2
351 (the most frequent CRC subtype, accounting for about 37% of cases) when compared to
352 normal tissue. In an unprecedented way, our results report the expression pattern of
353 these genes in consensus molecular subtypes of CRC. We also analyzed the expression
354 of the same genes of interest (*CLDN3*, and *N*-glycan-related genes) in the different
355 stages of CRC (Fig. 6, lower panel). Stage II tumors concomitantly exhibited low
356 expression levels of *CLDN3*, and *B3GNT8* when compared to their normal counterparts.
357 In addition, a downregulation of *B3GNT8* expression was also observed in all stages of
358 cancer. Interestingly, we observed that *MGAT5* was significantly upregulated in CRC,
359 even in the early stages of disease (I and II). This result suggests that the expression of
360 *MGAT5* could be considered a potential biomarker of CRC, once the gene expression
361 profile used in previous studies has been deployed to identify molecular biomarkers in
362 cancer (23).

363 **Discussion**

364 Stable apical junctional complex (AJC) has been considered as a suppressor of
365 carcinoma progression due to its role in the maintenance of apical–basolateral polarity,
366 intercellular adhesion, and epithelial architecture (34, 35). The dysregulation of this
367 protein complex is correlated with a malignant phenotype and a poor clinical outcome
368 (12, 36, 37). Although the regulatory role of *N*-glycans in the stability and function of
369 AJC has been demonstrated (16, 38), few studies have been dedicated to specifically
370 investigating the role of *N*-glycosylation on TJ function. Here, we have demonstrated
371 that the inhibition of the *N*-glycans biosynthesis pathway leads to claudin-3 subcellular
372 redistribution and decreases its levels in HCT-116 CRC cells (undifferentiated
373 phenotype). Changes in both protein levels (overexpression or downregulation) and
374 subcellular localization of different claudins may lead to the loss of TJ functionality (12,
375 39, 40). Previously, we reported that claudin-3 overexpression in HT-29 (moderately
376 differentiated CRC cells) increases the malignancy potential and affects the mechanisms
377 of paracellular flux control (12). In the present work, we observed that inhibition of
378 *N*-glycan biosynthesis by tunicamycin led to a decrease of claudin-3 levels in HCT-116
379 cells, which is a cell line that endogenously presents high levels of this protein,
380 however, tunicamycin did not fully affect the flow of ruthenium red dye through TJs.
381 We suspect that this finding may be related to the undifferentiated phenotype of these
382 cells, where the decrease in claudin-3 levels was not enough to completely restore
383 TJ-mediated permeability. Despite not determining full TJ function recovery, we
384 observed that tunicamycin led to the establishment of tighter cell–cell contacts and
385 promoted an increase of claudin-3 levels on the cellular membrane.

386 The influence of RTKs on the control of TJ stability is already known (13-15),
387 however, the regulatory role of *N*-glycans in this process remains poorly understood.
388 Here, we found that EGFR and IGF1R deglycosylation induced by treatment with
389 tunicamycin lead to a decrease in phosphorylation levels of both receptors. We also
390 demonstrated that the specific inhibition of EGFR or IGF1R decreased both their
391 phosphorylation levels and claudin-3 protein levels. These findings show that
392 RTK-related downstream signaling pathways regulate the content of claudin-3 in
393 HCT-116 cells. Although we have identified RTK-related signaling pathways (PLC and
394 STAT3) that regulate claudin-3 levels in colorectal cancer HCT-116 cells, this study was
395 deficient in demonstrating the mechanistic integration that interconnects RTKs,
396 glycogenes, and regulation of claudin-3 levels. Our ongoing studies aim to detail this
397 issue.

398 The differential levels of claudins in distinct carcinomas have been previously
399 reported (41). A gene expression-based study identified a molecular subtype of breast
400 cancer characterized by low levels of mRNA codifying for claudins, which was called
401 the claudin-low molecular subtype (42). Interestingly, while the low expression of
402 *CLDN3* in this subtype was related with worse overall survival (42), in other
403 carcinomas, such as colorectal, breast, gastric, ovary, and pancreas carcinomas, the
404 levels of various claudins were found to be increased (43), besides to being related with
405 a poor prognosis (44). Indeed, the expression levels of cancer-related genes were
406 extensively used as a parameter to determine tumor molecular subtypes related to
407 distinct clinical outcomes (45, 46). Regarding CRC, the recent gene profiling-based
408 stratification system, which has identified consensus molecular subtypes with
409 prognostic and predictive differences, represents a novel classification method to
410 improve clinical practice (47). In our study, we analyzed the overall survival of
411 colorectal cancer patients according to the expression levels of *CLDN3*. Low expression
412 levels of *CLDN3* in CMS2 and CMS3 improved the patients' long-term survival. Very
413 recently, similar results were reported regarding identification of intra-CMS subgroups
414 using the expression levels of claudins, thus corroborating the use of this strategy to
415 identify molecular subtypes (48).

416 Our findings revealed also that an integrated analysis of functionally related
417 genes in a particular cellular event (e.g., TJ stability regulation) should be explored as a
418 useful tool to better understand the specific alterations in each of the CMS.

419 Moreover, we also found that this supposed functional relation between *CLDN3*,
420 *MGAT5*, *ST6GAL1*, and *B3GNT8* could be subtype-specific, since we observed that
421 these genes display high expression levels in CMS2 as compared to CMS1. These
422 findings also suggest that multidimensional analyses considering not only different
423 stages of CRC but especially molecular subtypes are crucial for the identification of
424 regulatory mechanisms that rely on the integrated participation of several genes.

425 Although it can be assumed that the differential gene expression seen among
426 these CMS may have biological significance corresponding to their respective protein
427 levels, one cannot disregard the existence of discrepancies between mRNA levels and
428 protein expression attributable to other levels of regulation (49). Nevertheless, it has
429 been demonstrated that this correlation is significantly more reliable within distinct
430 biological groups (50).

431 An important issue in the classification of cancers in molecular subtypes
432 concerns the limitations imposed by tumor heterogeneity. Intratumoral heterogeneity

433 has challenged the actual classification of CRC because the region of the tumor where
434 the sample is taken for molecular profiling analysis could interfere in tumor
435 classification (51). Nevertheless, other authors have argued that tumor-intrinsic
436 subtyping captures the vast majority of biological diversity (52). In a pioneering work
437 regarding translational research in colorectal cancer, it was shown that a specific
438 molecular subtype of CRC, called CCS3, is resistant to anti-EGFR therapy in a clinical
439 setting independent of RAS mutation status, a classical determinant for therapy response
440 (53). Encouraging data have recently clarified this issue, showing that no differences
441 could be observed in the survival of KRAS/BRAF wild type patients treated with
442 cetuximab whose tumors had been classified as mesenchymal-like (CMS4) (54).
443 Therefore, in light of the discussion above, the importance of discovering both
444 molecular identities and new targets to improve the efficacy of therapies against CRC is
445 evident. Here, we identified an intra-CMS2 and intra-CMS3 subgroups with significant
446 differences for patients' long-term survival based on the expression pattern of *CLDN3*.
447 Furthermore, we demonstrated that the inhibition of *N*-glycan biosynthesis compromises
448 RTKs activation, thus corroborating previous data suggesting *N*-glycosylation as a
449 promising target in cancer therapy (55-57).

450 Aberrant *N*-glycosylation in cancer cells has been reported (58) and is regulated
451 among other factors by changes in enzyme levels that make up the glycosylation
452 machinery (46, 59). It is accepted that alterations in glycan structures as well as in the
453 enzymes responsible for them could be biomarkers in cancer (60-62). For example,
454 *MGAT5* and β 1,6-branched *N*-glycans are known to represent useful markers for
455 predicting the aggressive phenotype in CRC tumors (63, 64). Here, we observed that
456 *MGAT5* was upregulated in colorectal cancer samples, even in those belonging to early
457 stages of disease. This result suggests that the overexpression of this gene could be
458 considered as a potential CRC biomarker, once the *N*-glycan-related gene expression
459 profile has also been used to identify molecular biomarkers in cancer (46).

460 In conclusion, the data we have presented here show the modulatory role of
461 *N*-glycosylation on RTKs functionality and on regulation of claudin-3 protein levels. We
462 also demonstrated that the expression analysis of *CLDN3* and *N*-glycan-related genes
463 could be useful not only in determining clinically relevant colorectal cancer subtypes
464 but also in identifying potential glycobiomarkers. Moreover, our findings contribute to
465 explain how the dysregulation of claudin-3 in CRC occurs.

466 **Acknowledgment**

467 We are grateful to all members of the laboratory, particularly to Annie Cristhine Moraes
468 Sousa-Squiavinato and Bruna dos Santos Mendonça, for assistance with relevant
469 suggestions.

470 **Funding**

471 This work was funded by Ministério da Saúde, Coordenação de Aperfeiçoamento de
472 Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado do Rio
473 de Janeiro (FAPERJ), and Conselho Nacional de Desenvolvimento Científico e
474 Tecnológico (CNPq - grant number: 404052/2016-9).

475 **Availability of data and materials**

476 The data that support the findings of this study are available from the corresponding
477 author upon reasonable request.

478 **Authors' contributions**

479 This study was designed and supervised by JCM. AGP performed experiments and
480 drafted the manuscript. JAC, WFS, MSF, and CAF performed experiments. MB, CAL,
481 and PTS contributed in collecting and analyzing data. IMO and PVF collected patient
482 samples and performed experiments. JAM co-supervised the work.

483 **Ethics approval and consent to participate**

484 The study was approved by the Ethics Committee of the National Cancer Institute.
485 Written informed consent was obtained from all patients.

486 **Patient consent for publication**

487 Not applicable.

488 **Competing interests**

489 The authors have declared that no competing interest exists.

490 **Figure legends**

491 **Figure 1. Claudin-3 protein levels in CRC samples and survival features of**
492 **molecular subtypes substratified by *CLDN3* expression levels.** (A,B) Samples of
493 adjacent normal tissue (N) and tumor tissue (T) were obtained and processed for further
494 analysis of claudin-3 levels by Western blot ($n = 14$). Results of four representative
495 patients are shown. The line graph represents the increase (blue) or decrease (yellow) in
496 the ratio of claudin-3 densitometry units normalized by endogenous protein control
497 (GAPDH) in tumor and normal samples of each patient. (C, D, E, F, G) Prognostic
498 value of *CLDN3* expression (upper and lower tertiles) in unclassified data ($n = 414$) and
499 within CMS1 ($n = 69$), CMS2 ($n = 250$), CMS3 ($n = 44$), and CMS4 ($n = 52$).

500 **Figure 2: Effects of *N*-glycan biosynthesis inhibition on TJ stability.** (A) Cell lysates
501 from Caco-2, HT-29 and HCT-116 cells were analyzed by Western blot for claudin-3
502 (B) HCT-116 cells were treated with different concentrations of tunicamycin for 24 h.
503 After treatment, the cells were incubated with FITC-conjugated lectin L-PHA and
504 analyzed by flow cytometry. The histograms of fluorescence were generated by the Cell
505 Quest software: control (purple); 0.25 $\mu\text{g}/\text{mL}$ (green); 0.5 $\mu\text{g}/\text{mL}$ (pink); 0.75 $\mu\text{g}/\text{mL}$
506 (blue); and 1 $\mu\text{g}/\text{mL}$ (orange). (C) Lectin L-PHA specificity. (D) Cell lysates were
507 obtained after 24 h treatment with tunicamycin and analyzed by Western blot for
508 claudin-3, occludin, and ZO-1. Tubulin was used as an endogenous protein control. (E,
509 upper panel) Cell monolayers were fixed and stained for claudin-3 (green) and nucleus
510 (blue) (DAPI). Representative images were obtained by confocal microscopy. The
511 graphs represent the fluorescence intensity in cytosolic and membrane regions of
512 neighboring cells. Bar = 10 μm . (E, lower panel) Cells were cultured in filters of
513 Transwell polycarbonate, and the functionality of TJs was analyzed by transmission
514 electron microscopy (MET) using the ruthenium red tracer. The images are
515 representative of ultrathin sections of treated and control cells. Black arrows indicate the
516 cell-cell contact. Bar = 2 μm . The numerical values represent densitometric units \pm
517 standard error ($n = 3$). ns $P > 0.05$; * $P < 0.05$; ** $P < 0.01$; ANOVA. Tun, tunicamycin.

518 **Figure 3. Effects of *N*-glycan biosynthesis inhibition on RTKs functionality.** (A)
519 After treatment with tunicamycin, cell lysates were obtained and analyzed by Western
520 blot for p-EGFR, EGFR, p-IGF1R, and IGF1R. (B,C) Cells were treated with different
521 concentrations of PD153035 or OSI906 for 24 h, and then levels of RTK
522 phosphorylation and claudin-3 were assessed by Western blot. (D) Illustration showing
523 RTK-related or non-RTK-related pathways, as well as on the relationship of this
524 regulatory network with the regulation of *CLDN3* expression. **Effects of specific**
525 **inhibitors on claudin-3 levels.** (E, F, G, H) Cells were treated with different
526 concentrations of U73122, PD98059, Ly294002, and STA-21 for 24 h, and then levels
527 of claudin-3 were assessed by Western blot. (I) Cells were treated with combinations of
528 inhibitors: PD98059 (21 $\mu\text{g}/\text{mL}$) and Ly294002 (6 $\mu\text{g}/\text{mL}$), or STA-21 (7.5 $\mu\text{g}/\text{mL}$) and
529 U73122 (21 $\mu\text{g}/\text{mL}$). (J) Cells were treated with Forskolin or H-89 for 24 h, and then
530 levels of claudin-3 were assessed by Western blot. The numerical values represent
531 densitometric units \pm standard error ($n = 3$). ns $P > 0.05$; * $P < 0.05$; ** $P < 0.01$;
532 ANOVA. Tun, tunicamycin.

533 **Figure 4. Correlation analyses between *CLDN3* and *MGAT5*, *ST6GAL1*, and**
534 ***B3GNT8* within the CMS.** The correlation graphs represent the relationship between
535 gene expression data from the TCGA project database that were later classified for
536 stratification in the CMS.

537 **Figure 5. Correlation analyses between *CLDN3* and *MGAT5*, *ST6GAL1*, and**
538 ***B3GNT8* in each of the CRC stages.** The correlation graphs represent the relationship
539 between gene expression data from the TCGA project database that were stratified
540 based on the staging of the disease.

541 **Figure 6. Expression of *CLDN3*, and *N*-glycan-related genes (*MGAT5*, *ST6GAL1*,**
542 **and *B3GNT8*) in samples of CRC patients.** Box graphs represents absolute values of
543 gene expression from tumors ($n = 644$) and normal tissue samples ($n = 51$) accessed in
544 TCGA. (A) colorectal cancer consensus molecular subtypes (CMS). CMS1 $n = 90$;
545 CMS2 $n = 242$; CMS3 $n = 78$; CMS4 $n = 165$; and (B) colorectal cancer disease stages
546 (i, ii, iii, iv). ns $P > 0.05$; ** $P < 0.01$; *** $P < 0.001$. Yellow arrow indicates the

547 subtype with the lowest expression level of the analyzed gene; blue arrow indicates the
548 subtype with the highest expression level of the analyzed gene. Asterisks located above
549 box plots refer to comparisons made between each of the subtypes, and asterisks located
550 below box plots refer to comparisons made with respect to normal samples.

551 **Supplementary material**

552 **Table SI. Clinicopathological features of patients with colorectal cancer who were**
553 **analyzed in this study.** ¹Values are mean \pm standard deviation. ²Well, well-differentiated
554 adenocarcinoma; moderately, moderately differentiated adenocarcinoma with or without
555 mucinous areas; or undifferentiated adenocarcinoma; mucinous, mucinous
556 adenocarcinoma.

557 **Figure S1. N-glycan related enzymes.** The three glycosyltransferases (which encode
558 glycosyltransferases) analyzed in the study were *MGAT5*, *ST6GAL1*, and *B3GNT8*.
559 N-acetylglucosaminyltransferase V (*MGAT5*) catalyzes the transfer of GlcNAc in a β 1,6
560 linkage, generating branched *N*-glycans. *ST6GAL1* is a sialyltransferase that adds a
561 terminal α 2-6-linked sialic acid to the *N*-glycan. *B3GNT8* is an enzyme that is involved
562 in the biosynthesis of poly-N-acetyllactosamine chains by transferring GlcNAc to the
563 nonreducing terminus of Gal β 1-4GlcNAc on β 1,6-branched *N*-glycan.

564 **References**

- 565 1. IJspeert JE, Vermeulen L, Meijer GA and Dekker E: Serrated neoplasia-role in colorectal
566 carcinogenesis and clinical implications. *Nat Rev Gastroenterol Hepatol* 12: 401-409, 2015.
- 567 2. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA and Kinzler KW: Cancer
568 genome landscapes. *Science* 339: 1546-1558, 2013.

- 569 3. Fessler E, Drost J, van Hooff SR, et al.: TGF β signaling directs serrated adenomas to the
570 mesenchymal colorectal cancer subtype. *EMBO Mol Med* 8: 745-760, 2016.
- 571 4. Guinney J, Dienstmann R, Wang X, et al.: The consensus molecular subtypes of
572 colorectal cancer. *Nat Med* 21: 1350-1356, 2015.
- 573 5. Bramsen JB, Rasmussen MH, Ongen H, et al.: Molecular-Subtype-Specific Biomarkers
574 Improve Prediction of Prognosis in Colorectal Cancer. *Cell Rep* 19: 1268-1280, 2017.
- 575 6. Gehren AS, Rocha MR, de Souza WF and Morgado-Díaz JA: Alterations of the apical
576 junctional complex and actin cytoskeleton and their role in colorectal cancer progression.
577 *Tissue Barriers* 3: e1017688, 2015.
- 578 7. Lingaraju A, Long TM, Wang Y, Austin JR and Turner JR: Conceptual barriers to
579 understanding physical barriers. *Semin Cell Dev Biol* 42: 13-21, 2015.
- 580 8. Zihni C, Mills C, Matter K and Balda MS: Tight junctions: from simple barriers to
581 multifunctional molecular gates. *Nat Rev Mol Cell Biol* 17: 564-580, 2016.
- 582 9. de Oliveira SS, de Oliveira IM, De Souza W and Morgado-Díaz JA: Claudins upregulation
583 in human colorectal cancer. *FEBS Lett* 579: 6179-6185, 2005.
- 584 10. Krug SM, Schulzke JD and Fromm M: Tight junction, selective permeability, and related
585 diseases. *Semin Cell Dev Biol* 36: 166-176, 2014.
- 586 11. Wang Y, Sun T, Sun H, Yang S, Li D and Zhou D: SCF/C-Kit/JNK/AP-1 Signaling Pathway
587 Promotes Claudin-3 Expression in Colonic Epithelium and Colorectal Carcinoma. *Int J Mol Sci*
588 182017.
- 589 12. de Souza WF, Fortunato-Miranda N, Robbs BK, et al.: Claudin-3 Overexpression
590 Increases the Malignant Potential of Colorectal Cancer Cells: Roles of ERK1/2 and PI3K-Akt as
591 Modulators of EGFR signaling. *Plos One* 82013.
- 592 13. Singh AB and Harris RC: Epidermal growth factor receptor activation differentially
593 regulates claudin expression and enhances transepithelial resistance in Madin-Darby canine
594 kidney cells. *J Biol Chem* 279: 3543-3552, 2004.
- 595 14. Ikari A, Sato T, Watanabe R, Yamazaki Y and Sugatani J: Increase in claudin-2 expression
596 by an EGFR/MEK/ERK/c-Fos pathway in lung adenocarcinoma A549 cells. *Biochim Biophys Acta*
597 1823: 1110-1118, 2012.
- 598 15. Campbell HK, Maiers JL and DeMali KA: Interplay between tight junctions & adherens
599 junctions. *Exp Cell Res* 2017.
- 600 16. Nita-Lazar M, Rebutini I, Walker J and Kukuruzinska MA: Hypoglycosylated E-cadherin
601 promotes the assembly of tight junctions through the recruitment of PP2A to adherens
602 junctions. *Exp Cell Res* 316: 1871-1884, 2010.
- 603 17. Fernandes H, Cohen S and Bishayee S: Glycosylation-induced conformational
604 modification positively regulates receptor-receptor association: a study with an aberrant
605 epidermal growth factor receptor (EGFRvIII/DeltaEGFR) expressed in cancer cells. *J Biol Chem*
606 276: 5375-5383, 2001.
- 607 18. Whitson KB, Whitson SR, Red-Brewer ML, et al.: Functional effects of glycosylation at
608 Asn-579 of the epidermal growth factor receptor. *Biochemistry* 44: 14920-14931, 2005.
- 609 19. Kaszuba K, Grzybek M, Orłowski A, et al.: N-Glycosylation as determinant of epidermal
610 growth factor receptor conformation in membranes. *Proc Natl Acad Sci U S A* 112: 4334-4339,
611 2015.
- 612 20. Lopez Sambrooks C, Baro M, Quijano A, et al.: Oligosaccharyltransferase Inhibition
613 Overcomes Therapeutic Resistance to EGFR Tyrosine Kinase Inhibitors. *Cancer Res* 78:
614 5094-5106, 2018.
- 615 21. Lau KS, Partridge EA, Grigorian A, et al.: Complex N-glycan number and degree of
616 branching cooperate to regulate cell proliferation and differentiation. *Cell* 129: 123-134, 2007.
- 617 22. Colaprico A, Silva TC, Olsen C, et al.: TCGAAbiolinks: an R/Bioconductor package for
618 integrative analysis of TCGA data. *Nucleic Acids Res* 44: e71, 2016.
- 619 23. Fekete T, Rásó E, Pete I, et al.: Meta-analysis of gene expression profiles associated with
620 histological classification and survival in 829 ovarian cancer samples. *Int J Cancer* 131: 95-105,
621 2012.

- 622 24. Motamed-Khorasani A, Jurisica I, Letarte M, et al.: Differentially androgen-modulated
623 genes in ovarian epithelial cells from BRCA mutation carriers and control patients predict
624 ovarian cancer survival and disease progression. *Oncogene* 26: 198-214, 2007.
- 625 25. de Freitas Junior JC, Silva BuR, de Souza WF, de Araújo WM, Abdelhay ES and
626 Morgado-Díaz JA: Inhibition of N-linked glycosylation by tunicamycin induces
627 E-cadherin-mediated cell-cell adhesion and inhibits cell proliferation in undifferentiated human
628 colon cancer cells. *Cancer Chemother Pharmacol* 68: 227-238, 2011.
- 629 26. Nami B, Donmez H and Kocak N: Tunicamycin-induced endoplasmic reticulum stress
630 reduces in vitro subpopulation and invasion of CD44+/CD24- phenotype breast cancer stem
631 cells. *Exp Toxicol Pathol* 68: 419-426, 2016.
- 632 27. Peter Y, Comellas A, Levantini E, Ingenito EP and Shapiro SD: Epidermal growth factor
633 receptor and claudin-2 participate in A549 permeability and remodeling: implications for
634 non-small cell lung cancer tumor colonization. *Mol Carcinog* 48: 488-497, 2009.
- 635 28. Taniguchi N and Korekane H: Branched N-glycans and their implications for cell
636 adhesion, signaling and clinical applications for cancer biomarkers and in therapeutics. *BMB*
637 *Rep* 44: 772-781, 2011.
- 638 29. Ishida H, Togayachi A, Sakai T, et al.: A novel beta1,3-N-acetylglucosaminyltransferase
639 (beta3Gn-T8), which synthesizes poly-N-acetyllactosamine, is dramatically upregulated in colon
640 cancer. *FEBS Lett* 579: 71-78, 2005.
- 641 30. Dall'Olio F, Malagolini N, di Stefano G, Minni F, Marrano D and Serafini-Cessi F:
642 Increased CMP-NeuAc:Gal beta 1,4GlcNAc-R alpha 2,6 sialyltransferase activity in human
643 colorectal cancer tissues. *Int J Cancer* 44: 434-439, 1989.
- 644 31. Petretti T, Kemmner W, Schulze B and Schlag PM: Altered mRNA expression of
645 glycosyltransferases in human colorectal carcinomas and liver metastases. *Gut* 46: 359-366,
646 2000.
- 647 32. Dall'Olio F, Chiricolo M, Ceccarelli C, Minni F, Marrano D and Santini D: Beta-galactoside
648 alpha2,6 sialyltransferase in human colon cancer: contribution of multiple transcripts to
649 regulation of enzyme activity and reactivity with Sambucus nigra agglutinin. *Int J Cancer* 88:
650 58-65, 2000.
- 651 33. Dall'Olio F, Chiricolo M, Mariani E and Facchini A: Biosynthesis of the cancer-related
652 sialyl-alpha 2,6-lactosaminyl epitope in colon cancer cell lines expressing beta-galactoside
653 alpha 2,6-sialyltransferase under a constitutive promoter. *Eur J Biochem* 268: 5876-5884, 2001.
- 654 34. Royer C and Lu X: Epithelial cell polarity: a major gatekeeper against cancer? *Cell Death*
655 *Differ* 18: 1470-1477, 2011.
- 656 35. Halaoui R and McCaffrey L: Rewiring cell polarity signaling in cancer. *Oncogene* 34:
657 939-950, 2015.
- 658 36. Bouchagier KA, Assimakopoulos SF, Karavias DD, et al.: Expression of claudins-1, -4, -5,
659 -7 and occludin in hepatocellular carcinoma and their relation with classic clinicopathological
660 features and patients' survival. *In Vivo* 28: 315-326, 2014.
- 661 37. Katayama A, Handa T, Komatsu K, et al.: Expression patterns of claudins in patients with
662 triple-negative breast cancer are associated with nodal metastasis and worse outcome. *Pathol*
663 *Int* 67: 404-413, 2017.
- 664 38. Pinho SS, Figueiredo J, Cabral J, et al.: E-cadherin and adherens-junctions stability in
665 gastric carcinoma: functional implications of glycosyltransferases involving N-glycan branching
666 biosynthesis, N-acetylglucosaminyltransferases III and V. *Biochim Biophys Acta* 1830:
667 2690-2700, 2013.
- 668 39. Bücken R, Krug SM, Fromm A, et al.: Campylobacter fetus impairs barrier function in
669 HT-29/B6 cells through focal tight junction alterations and leaks. *Ann N Y Acad Sci* 2017.
- 670 40. Zeissig S, Bürgel N, Günzel D, et al.: Changes in expression and distribution of claudin 2,
671 5 and 8 lead to discontinuous tight junctions and barrier dysfunction in active Crohn's disease.
672 *Gut* 56: 61-72, 2007.
- 673 41. Kwon MJ: Emerging roles of claudins in human cancer. *Int J Mol Sci* 14: 18148-18180,
674 2013.

- 675 42. Dias K, Dvorkin-Gheva A, Hallett RM, et al.: Claudin-Low Breast Cancer; Clinical &
676 Pathological Characteristics. *PLoS One* 12: e0168669, 2017.
- 677 43. Oliveira SS and Morgado-Díaz JA: Claudins: multifunctional players in epithelial tight
678 junctions and their role in cancer. *Cell Mol Life Sci* 64: 17-28, 2007.
- 679 44. Iacobuzio-Donahue CA, Maitra A, Shen-Ong GL, et al.: Discovery of novel tumor
680 markers of pancreatic cancer using global gene expression technology. *Am J Pathol* 160:
681 1239-1249, 2002.
- 682 45. Levine EA, Votanopoulos KI, Qasem SA, et al.: Prognostic Molecular Subtypes of
683 Low-Grade Cancer of the Appendix. *J Am Coll Surg* 222: 493-503, 2016.
- 684 46. Ashkani J and Naidoo KJ: Glycosyltransferase Gene Expression Profiles Classify Cancer
685 Types and Propose Prognostic Subtypes. *Sci Rep* 6: 26451, 2016.
- 686 47. Thanki K, Nicholls ME, Gajjar A, et al.: Consensus Molecular Subtypes of Colorectal
687 Cancer and their Clinical Implications. *Int Biol Biomed J* 3: 105-111, 2017.
- 688 48. Cherradi S, Martineau P, Gongora C and Del Rio M: Claudin gene expression profiles
689 and clinical value in colorectal tumors classified according to their molecular subtype. *Cancer*
690 *Manag Res* 11: 1337-1348, 2019.
- 691 49. Vogel C and Marcotte EM: Insights into the regulation of protein abundance from
692 proteomic and transcriptomic analyses. *Nat Rev Genet* 13: 227-232, 2012.
- 693 50. Koussounadis A, Langdon SP, Um IH, Harrison DJ and Smith VA: Relationship between
694 differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. *Sci*
695 *Rep* 5: 10775, 2015.
- 696 51. Dunne PD, McArt DG, Bradley CA, et al.: Challenging the Cancer Molecular Stratification
697 Dogma: Intratumoral Heterogeneity Undermines Consensus Molecular Subtypes and Potential
698 Diagnostic Value in Colorectal Cancer. *Clin Cancer Res* 22: 4095-4104, 2016.
- 699 52. Prat A, Pineda E, Adamo B, et al.: Clinical implications of the intrinsic molecular
700 subtypes of breast cancer. *Breast* 24 Suppl 2: S26-35, 2015.
- 701 53. De Sousa E Melo F, Wang X, Jansen M, et al.: Poor-prognosis colon cancer is defined by
702 a molecularly distinct subtype and develops from serrated precursor lesions. *Nat Med* 19:
703 614-618, 2013.
- 704 54. Trinh A, Trumpi K, De Sousa E Melo F, et al.: Practical and Robust Identification of
705 Molecular Subtypes in Colorectal Cancer by Immunohistochemistry. *Clin Cancer Res* 23:
706 387-398, 2017.
- 707 55. de Freitas Junior JC and Morgado-Díaz JA: The role of N-glycans in colorectal cancer
708 progression: potential biomarkers and therapeutic applications. *Oncotarget* 7: 19395-19413,
709 2016.
- 710 56. Carvalho S, Reis CA and Pinho SS: Cadherins Glycans in Cancer: Sweet Players in a Bitter
711 Process. *Trends Cancer* 2: 519-531, 2016.
- 712 57. Contessa JN, Bhojani MS, Freeze HH, Ross BD, Rehemtulla A and Lawrence TS:
713 Molecular imaging of N-linked glycosylation suggests glycan biosynthesis is a novel target for
714 cancer therapy. *Clin Cancer Res* 16: 3205-3214, 2010.
- 715 58. Anugraham M, Jacob F, Nixdorf S, Everest-Dass AV, Heinzelmann-Schwarz V and Packer
716 NH: Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: glycan
717 structures reflect gene expression and DNA methylation status. *Mol Cell Proteomics* 13:
718 2213-2232, 2014.
- 719 59. Kannagi R, Yin J, Miyazaki K and Izawa M: Current relevance of incomplete synthesis
720 and neo-synthesis for cancer-associated alteration of carbohydrate determinants--Hakomori's
721 concepts revisited. *Biochim Biophys Acta* 1780: 525-531, 2008.
- 722 60. Meany DL and Chan DW: Aberrant glycosylation associated with enzymes as cancer
723 biomarkers. *Clin Proteomics* 8: 7, 2011.
- 724 61. Liu L, Yan B, Huang J, et al.: The identification and characterization of novel
725 N-glycan-based biomarkers in gastric cancer. *PLoS One* 8: e77821, 2013.
- 726 62. Qin R, Zhao J, Qin W, et al.: Discovery of Non-invasive Glycan Biomarkers for Detection
727 and Surveillance of Gastric Cancer. *J Cancer* 8: 1908-1916, 2017.

- 728 63. Kim YS, Ahn YH, Song KJ, et al.: Overexpression and
729 β -1,6-N-acetylglucosaminylation-initiated aberrant glycosylation of TIMP-1: a "double
730 whammy" strategy in colon cancer progression. *J Biol Chem* 287: 32467-32478, 2012.
- 731 64. Murata K, Miyoshi E, Kameyama M, et al.: Expression of
732 N-acetylglucosaminyltransferase V in colorectal cancer correlates with metastasis and poor
733 prognosis. *Clin Cancer Res* 6: 1772-1777, 2000.

Patients	14
CLINICOPATHOLOGICAL FEATURES	
Gender	
Male/Female	5/9
Age (years)¹	64± 10
Location	
Terminal ileum and right colon	4
Right Colon	4
Left colon	1
Sigmoid colon	5
Histology²	
Well	1
Moderately	11
Mucinous	2
TNM stage	
0	1
I	2
II	5
III	6

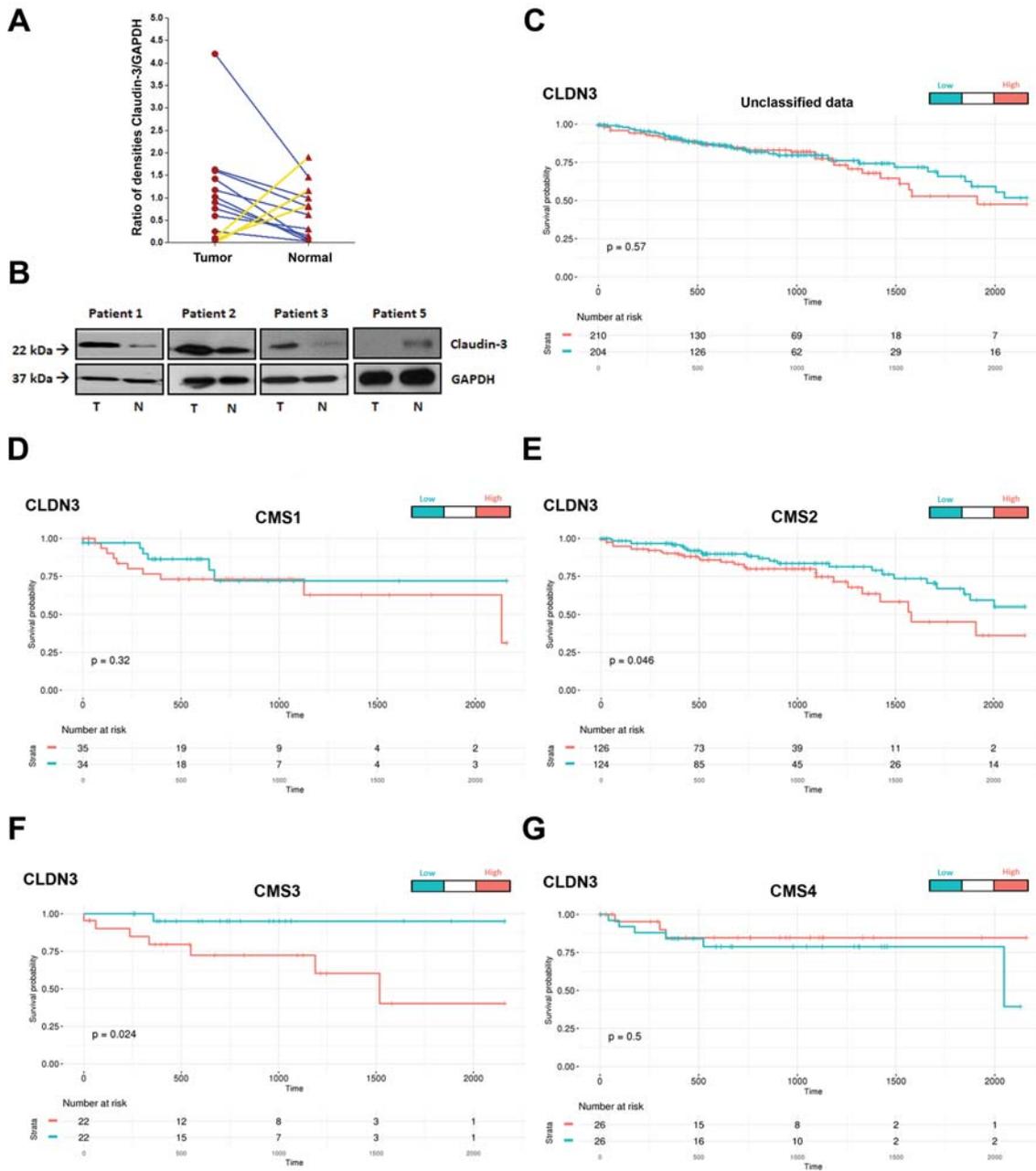


Figure 01

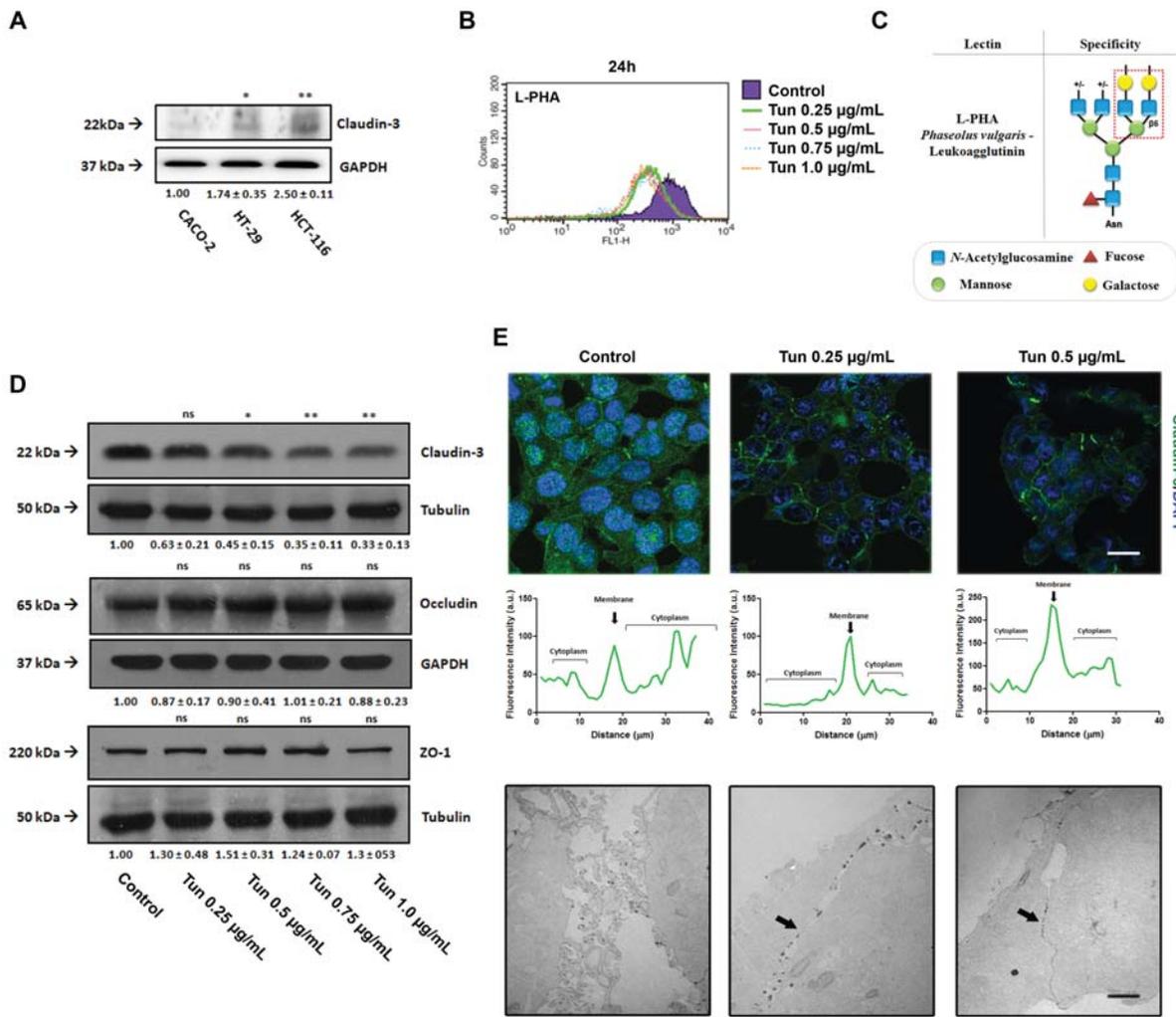


Figure 02

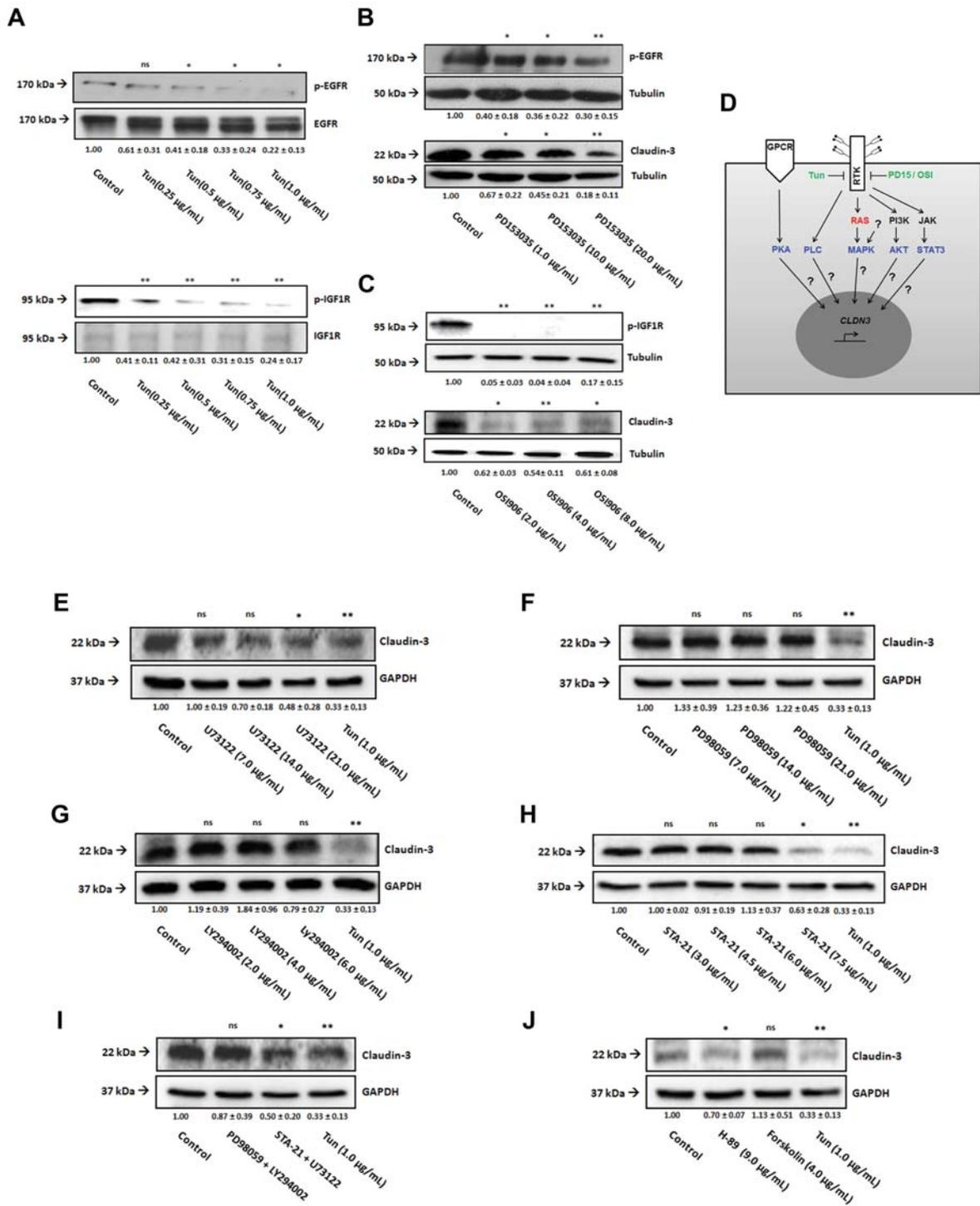
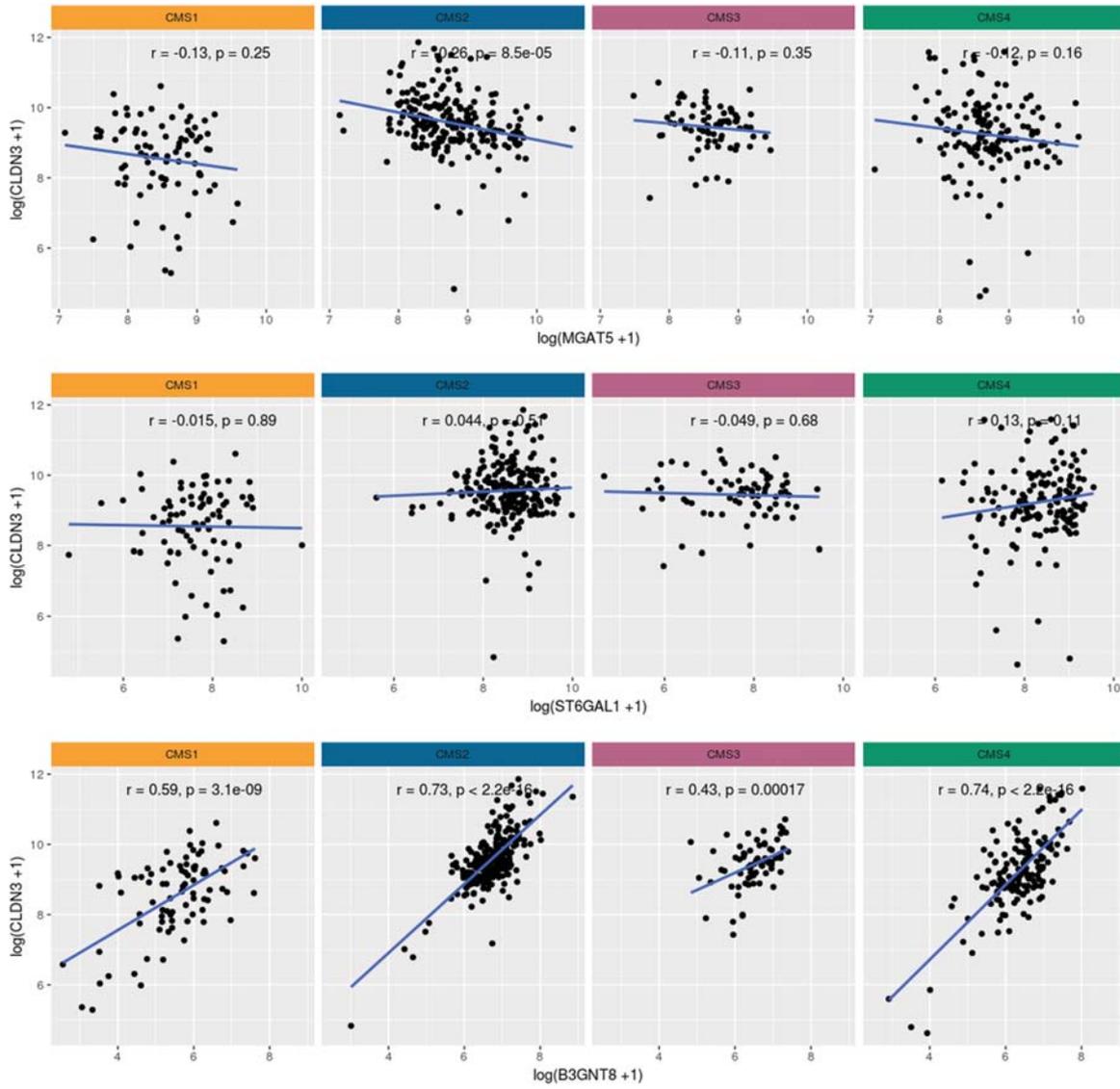


Figure 03

**Figure 04**

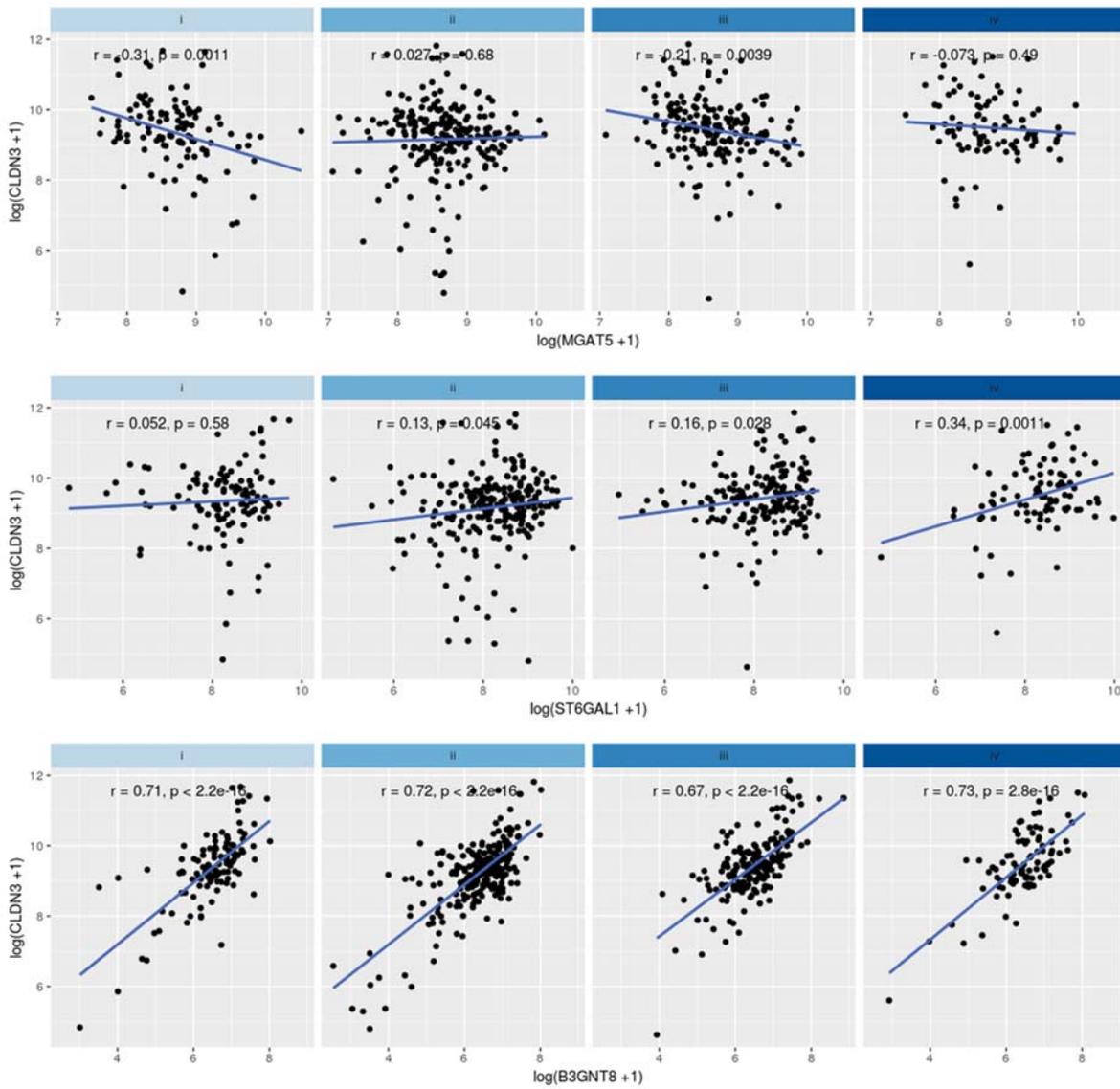


Figure 05

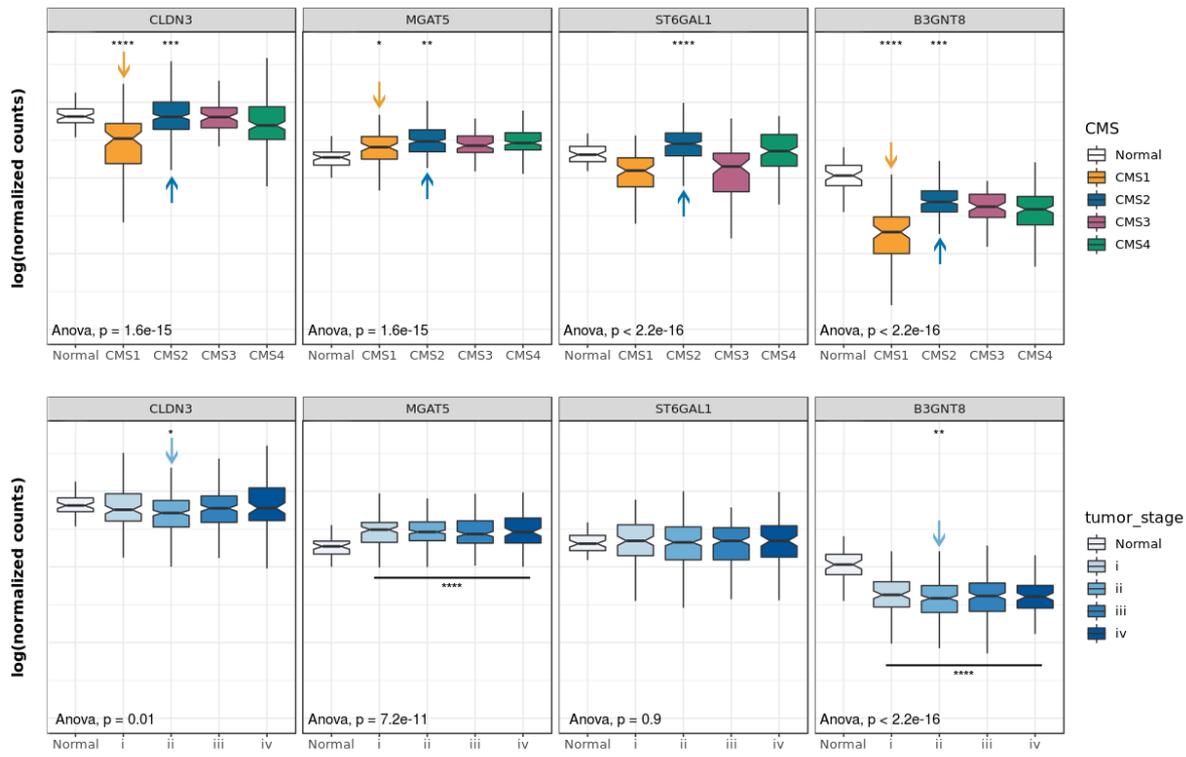
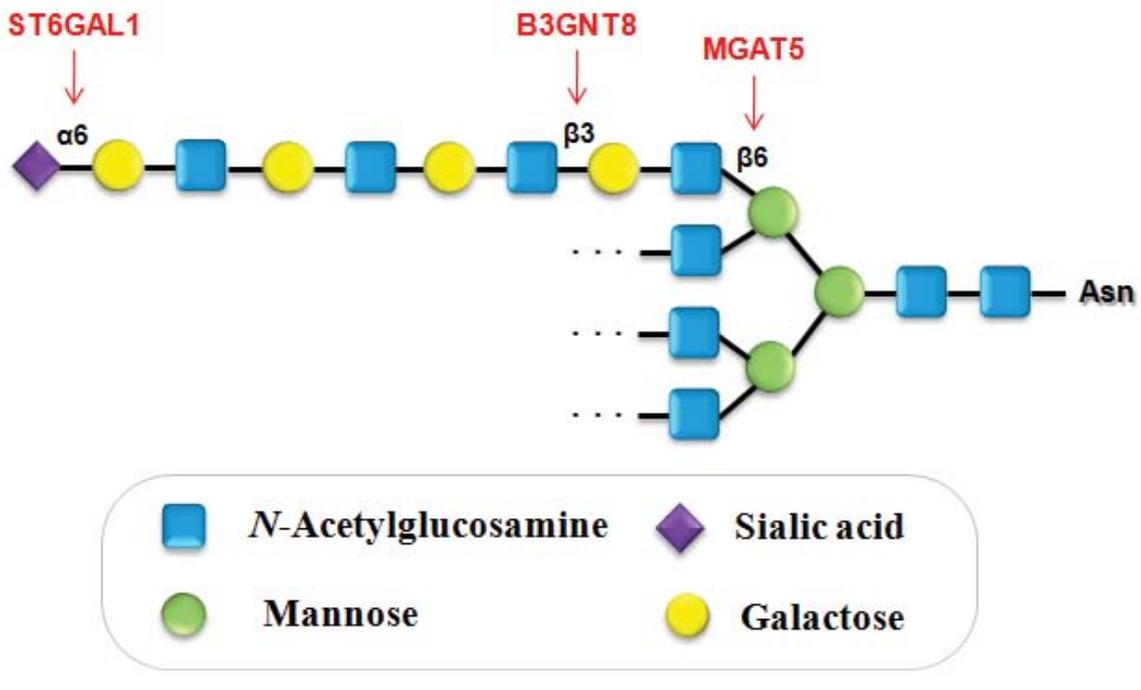


Figure 06



Supplementary material - Figure S1

Fw: Successful Submission of a Manuscript to Science Translational Medicine (abb5958)

Julio de Freitas <juliofreitasjr@yahoo.com.br>

Sex, 13/03/2020 14:17

Para: Jessica Andrade <jessica_andradecos@hotmail.com>

Manuscript Title: Glycans as immune-checkpoints: removal of branched N-glycans enhances immune recognition associate with cancer prevention.

Author: de-Freitas-Junior

Manuscript Number: abb5958

Dear Dr. de-Freitas-Junior

You are listed as a coauthor on the above manuscript, which has recently been submitted to Science Translational Medicine. According to Science Translational Medicine policy, all authors must have seen and approved the submission of their manuscript. If you have seen the manuscript and approved its submission, no action is necessary.

If you have not read this paper or do not approve its submission to Science Translational Medicine, please let us know as soon as possible. Please refer to the manuscript number listed above in any correspondence (you can just reply to this message).

You can see the status of your manuscript at any time by logging into your account at the Science Journals Manuscript Submission and Information Portal at <https://cts.sciencemag.org>.

If you do not yet have an account, go to the website and create one with the email address at which you received this correspondence, as this email address is linked to your manuscript. You must then send an email to scitranslmededitors@aaas.org to request that your active manuscript be added to your account. We will confirm activation and you can then login, access your manuscript and monitor its status.

Once you have an account, we also encourage you to link your account to your ORCID ID, an identifier that facilitates the correct attribution of your publications to you. To learn more about ORCID or to obtain an ORCID ID, visit their site at: <http://orcid.org>.

Your manuscript is now undergoing an initial screening to determine whether it will be sent for in-depth review.

Sincerely,

The Editors
Science Translational Medicine