

**IDENTIFICAÇÃO DE MODULADORES DA
SINALIZAÇÃO CELULAR NO CÂNCER DE CÓLON E
METÁSTASES HEPÁTICAS COMO FERRAMENTA
PARA DESCOBERTA DE ALVOS TERAPÊUTICOS**

DANIELA BIZINELLI

**Dissertação apresentada à Fundação Antônio Prudente para
obtenção de Título de Mestre em Ciências**

Área de concentração: Oncologia

Orientador: Dr. Fabio Albuquerque Marchi

São Paulo

2022

FICHA CATALOGRÁFICA

Bizinelli, Daniela.

Identificação de moduladores da sinalização celular no câncer de cólon e metástases hepáticas como ferramenta para descoberta de alvos terapêuticos

/ Daniela Bizinelli - São Paulo, 2022.

120p.

Dissertação (Mestrado)-Fundação Antônio Prudente.

Curso de Pós-Graduação em Ciências - Área de concentração: Oncologia.

Orientador: Fabio Albuquerque Marchi.

Descritores. 1. Neoplasias colônicas/Colonic Neoplasms. 2. Metástase Neoplásica/Neoplasm Metastasis. 3. Proteômica/Proteomics. 4 Resistência a Medicamentos Antineoplásicos/Drug Resistance, Neoplasm. 5. Transdução de Sinais/Signal Transduction. 6. Prognóstico/Prognosis.

DANIELA BIZINELLI

**IDENTIFICAÇÃO DE MODULADORES DA SINALIZAÇÃO CELULAR
NO CÂNCER DE CÓLON E METÁSTASES HEPÁTICAS COMO
FERRAMENTA PARA DESCOBERTA DE ALVOS TERAPÊUTICOS**

Aprovado em: 06/09/2022

Banca examinadora

Orientador: Dr. Fabio Albuquerque Marchi

Instituição: Fundação Antônio Prudente (A.C.Camargo Cancer Center)

Membro da banca: Dr. Roberto César Pereira Lima Júnior

Instituição: Universidade Federal do Ceará (UFC)

Membro da banca: Dr. Ulysses Ribeiro Júnior

Instituição: Faculdade de Medicina da Universidade de São Paulo, Instituto do Câncer do Estado de São Paulo (ICESP-HCFMUSP)

Membro da banca: Dr. Virgilio Souza e Silva

Instituição: Fundação Antônio Prudente (A.C.Camargo Cancer Center)

SUPORTE À PESQUISA POR AGÊNCIA DE FOMENTO

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processo nº 133699/2020-0, e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo nº 2021/01298-0. As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade dos autores e não necessariamente refletem a visão do CNPq e da FAPESP.

*“O mar escuro trará o medo lado a lado
com os corais mais coloridos”.*

O Rappa

Dedico esse trabalho à minha família, que sempre esteve
presente durante toda a jornada.

Aos pacientes, que mesmo com tantas adversidades,
acreditam na ciência e em um futuro próspero.

AGRADECIMENTOS

Agradeço primeiramente aos meus pais, Daniel e Lilian, por terem sempre buscado oferecer mais do que a base para o meu desenvolvimento pessoal. Obrigada por nunca medirem esforços e por se fazerem tão presentes na minha vida. Obrigada por se interessarem, vibrarem e se preocuparem, mesmo quando vocês não entendem as entrelinhas. Vocês são meu exemplo. Obrigada por tanto.

À minha família e, principalmente, aos meus avós, que sempre desempenharam um papel fundamental em minha trajetória. Amo vocês.

Ao meu orientador, Dr. Fabio Albuquerque Marchi, pela confiança e dedicação a esse trabalho. Obrigada por ter me aceito como aluna, mesmo com pouco conhecimento prévio em bioinformática, e por tanto ter me ensinado, sempre disponível e paciente. Agradeço pelas oportunidades e pelo crescimento científico que pude ter durante esse período.

À todas as pessoas que passaram pela minha vida e compartilharam experiências durante esses dois anos. Em especial, agradeço à Anna, Julia, Ingrid, Danilo, Jacque, Grazi, Bruna, Fer e Bianca, que tornaram meus dias mais leves.

Ao Programa de Pós-Graduação do A.C.Camargo Cancer Center e aos funcionários do hospital e do CIPE. Obrigada por direta ou indiretamente nos auxiliarem durante o desenvolvimento de nossas pesquisas.

Aos membros da Banca de Qualificação e Defesa, pela disponibilidade, contribuições e auxílio no direcionamento deste projeto.

Ao *Core Facility* de Histologia e Imuno-histoquímica, em especial à Keila e Nathalia, pela disponibilidade e auxílio durante o resgate e preparação de lâminas e obtenção dos materiais emblocados em parafina.

Ao Departamento de Patologia, em especial ao Dr. Gabriel Oliveira dos Santos, pela disponibilidade e agilidade na avaliação das lâminas de H&E.

À Dra. Cláudia Malheiros, Dra. Milena Monteiro e Katia Klug, pelo auxílio nos protocolos, disponibilidade e atenção para sanar todas as minhas dúvidas.

Ao Laboratório Multiusuário de Proteômica, Metabolômica e Lipidômica Max Feffer, no Departamento de Genética na Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP), Dr. Carlos Alberto Labate e Dra. Mônica Teresa Veneziano Labate, pela recepção, auxílio nos experimentos e ajuda na coleta dos resultados.

Às minhas amigas, Beatriz, Juliana e Thais, pela cumplicidade e amizade de longa data. Sempre estarei aqui por vocês, assim como vocês sempre estiveram por mim.

Ao meu companheiro Guilherme, por entender minhas ausências, escutar minhas angústias e incentivar meus sonhos. É e sempre será um prazer partilhar a vida com você.

Por fim, agradeço ao CNPq e à FAPESP pelo incentivo à pesquisa e auxílio.

RESUMO

Bizinelli Daniela. **Identificação de moduladores da sinalização celular no câncer de cólon e metástases hepáticas como ferramenta para descoberta de alvos terapêuticos.** [Dissertação]. São Paulo; Fundação Antônio Prudente; 2022.

O câncer de cólon é uma das principais causas de morte por câncer em todo mundo e aproximadamente 50% dos pacientes desenvolvem metástases hepáticas durante o curso da doença. Ainda que a ressecção cirúrgica proporcione sobrevida de 5 anos ao redor de 40%, a maior parte dos pacientes apresenta doença irresssecável ao diagnóstico. Neste caso, o cenário é devastador devido à resistência das células tumorais ao tratamento padrão com quimioterapia e/ou anticorpo monoclonal. Apesar dos avanços alcançados pelo rastreamento genômico das metástases, o conhecimento sobre o envolvimento de transcritos e proteínas específicas na atividade de sinalização e recrutamento de populações imunes ainda é limitado nessa doença. Considerando o grande potencial de descoberta e inovação que pode ser alcançado, este projeto propôs o desenvolvimento de uma meta-análise de dados públicos para a exploração de metástases hepáticas e a caracterização proteômica de metástases hepáticas inicialmente irresssecáveis em busca de alvos promissores que possam ser explorados em novas opções terapêuticas. As principais etapas metodológicas incluem (1) desenvolvimento de um modelo de integração de dados para a obtenção, processamento e análise de aproximadamente 3.5 mil amostras de RNA-seq, *microarray* e proteômica de metástases hepáticas inicialmente ressecáveis, tumores primários do cólon e tecidos não-neoplásicos, (2) rastreamento por espectrometria de massas de 30 biópsias pareadas, sendo 15 de metástases hepáticas inicialmente irresssecáveis e 15 de tecidos não-neoplásicos adjacentes, (3) investigação de potenciais moduladores da sinalização oncogênica associada à resistência à terapia e (4) exploração dos resultados em um conjunto independente com aproximadamente 138 mil células obtidas por *single-cell* RNA-seq em busca de associações entre alvos, vias de sinalização e populações imunes. Após a análise de espectrometria de massas e integração de dados da meta-análise, das 46 proteínas diferencialmente reguladas entre metástases hepáticas inicialmente irresssecáveis e tecidos não-neoplásicos, TGFB1, CDH2, APOA1, TNFR2, EGFR, MDH2, ITIH2, YWHAZ, PCBP1, TOP2A e ENO1 (todas com aumento de regulação nas metástases) e CXCL14, PLA2G2A e CAV1 (com diminuição de regulação nas metástases) prosseguiram

na análise. Dessas, PLA2G2A e CAV1 não foram estatisticamente confirmadas após comparação com cinco estudos independentes. Também foi identificada a ativação nas metástases de importantes vias como JAK-STAT ($z\text{-score}=3,18$), VEGF ($z\text{-score}=2,67$), WNT ($z\text{-score}=3,22$), PI3K ($z\text{-score}=3,99$), MAPK ($z\text{-score}=2,85$), EGFR ($z\text{-score}=2,49$) e NFkB ($z\text{-score}=4,11$). Em seguida observamos pela análise de *single-cell* RNA-seq que a presença de *TGFB1*, *TNFR2* e *EGFR* estava associada a maior polarização de neutrófilos e células Tregs em metástases hepáticas, sugerindo uma participação desses alvos no recrutamento dessas populações imunes ou em conjunto durante o desenvolvimento, estabelecimento dessas células no fígado e progressão da doença. Tanto *TGFB1* ($p=0,01$) quanto *TNFR2* ($p=0,048$) foram associados a pior sobrevida global em pacientes metastáticos e apresentaram AUC=0,95 na separação entre câncer primário no cólon e metástases no fígado. Especificamente sobre a comparação entre metástases inicialmente ressecáveis e irressecáveis, apesar de termos identificado alguns alvos com mais de 10-fold de diferença (YWHAZ, AHNAK, HNRNPH1, PCBP1 e TGFB1) apenas TGFB1 prosseguiu em todas as análises, porém, esses resultados sugerem a exploração posterior desses candidatos em novos estudos. Quando buscamos por alvos relacionados à resistência à terapia, nossa estratégia selecionou 32 proteínas diferencialmente reguladas nas metástases resistentes, com envolvimento de CDH2 ($p=0,002$), CXCL14 ($p=0,05$), SERPINA1 ($p=0,012$) e LRGR ($p=0,049$) com pior sobrevida global em pacientes metastáticos. Nesses tumores também foi observada uma tendência de ativação das vias TGF- β e TNF- α , já descritas na literatura devido a participação no desenvolvimento e progressão de metástases, além de favorecer um perfil de resistência às células tumorais. Por meio de uma análise de deconvolução também observamos maior presença de células Tregs em pacientes resistentes. Essa relação entre Tregs/LGR5 também já foi associada a pior prognóstico em outros tipos de cânceres e suportam uma participação de ambos em um possível mecanismo de resistência desses tumores. Com base nos resultados obtidos, acreditamos oferecer um conjunto de dados inéditos que podem modificar o prognóstico e diagnóstico das metástases hepáticas, bem como contribuir para que os pacientes tenham novas opções terapêuticas com melhora na sobrevida.

Palavras-chave: metástase hepática; proteômica; espectrometria de massas; marcador prognóstico; resistência terapêutica.

ABSTRACT

Bizinelli Daniela. **[Identification of cell signaling modulators in colon cancer and liver metastases as a tool for discovering therapeutic targets]**. [Dissertação]. São Paulo; Fundação Antônio Prudente; 2022.

Colon cancer is one of the leading causes of cancer death worldwide and approximately 50% of patients develop liver metastases during the course of the disease. Although surgical resection provides a 5-year survival rate of around 40%, most patients have unresectable disease at diagnosis. In this case, the scenario is devastating due to the resistance of tumor cells to standard treatment with chemotherapy and/or monoclonal antibody. Despite the advances achieved by genomic tracking of metastases, the involvement of specific transcripts and proteins in the signaling activity and recruitment of immune populations is still limited in this disease. Considering the great potential for discovery and innovation that can be achieved, this project proposed the development of a meta-analysis for the exploration of liver metastases and the proteomic characterization of initially unresectable liver metastases in search of promising targets that can be explored in new therapeutics options. Key methodological steps include (1) development of a data integration model for obtaining, processing and analyzing approximately 3,500 RNA-seq, microarray and proteomic samples from initially resectable liver metastases, primary colon tumors and non-neoplastic tissues, (2) mass spectrometry screening of 30 paired biopsies, 15 from initially unresectable liver metastases and 15 from adjacent non-neoplastic tissues, (3) investigation of potential modulators of oncogenic signaling associated with therapy resistance, and (4) exploration of the results in an independent pool of approximately 138,000 cells obtained by single-cell RNA-seq in search of associations between targets, signaling pathways and immune populations. After mass spectrometry analysis and integration of meta-analysis data, of the 46 differentially regulated proteins between initially unresectable liver metastases and non-neoplastic tissues, TGFB1, CDH2, APOA1, TNFR2, EGFR, MDH2, ITIH2, YWHAZ, PCBP1, TOP2A and ENO1 (all with upregulation in metastases) and CXCL14, PLA2G2A and CAV1 (with downregulation in metastases) continued in the analysis. Of these, PLA2G2A and CAV1 were not statistically confirmed after comparison with five independent studies. We identified activation of important pathways such as JAK-STAT (z-score=3.18), VEGF (z-score=2.67), WNT (z-score=3.22), PI3K (z-score=3.99), MAPK (z-

score=2.85), EGFR (z-score=2.49) and NFkB (z-score=4.11). Then we observed by single-cell RNA-seq analysis that the presence of *TGFB1*, *TNFR2* and *EGFR* was associated with greater polarization of neutrophils and Treg cells in liver metastases, suggesting a participation of these targets in the recruitment of these immune populations or together during development, establishment of these cells in the liver and disease progression. Both *TGFB1* (p=0.01) and *TNFR2* (p=0.048) were associated with worse overall survival in patients with metastatic disease and had AUC=0.95 separating primary colon cancer from liver metastases. Specifically, regarding the comparison between initially resectable and unresectable metastases, although we identified some targets with more than 10-fold difference (YWHAZ, AHNAK, HNRNPH1, PCBP1 and TGFB1) only TGFB1 continued in all analyses, however, these results suggest the exploration of these candidates in further studies. When looking for targets related to therapy resistance, our strategy selected 32 proteins differentially regulated in resistant metastases, with involvement of CDH2 (p=0.002), CXCL14 (p=0.05), SERPINA1 (p=0.012) and LRGR (p=0.049) with poor overall survival in metastatic patients. In these tumors, a tendency of activation of the TGF- β and TNF- α pathways was also observed, already described in the literature due to their participation in the development and progression of metastases, in addition to favoring a resistance profile to tumor cells. By means of a deconvolution analysis, we also observed a greater presence of Treg cells in resistant patients. This relationship between Tregs/LGR5 has also been associated with a worse prognosis in other types of cancers and supports the participation of both in a possible mechanism of resistance in these tumors. Based on the results obtained, we believe to offer a set of unpublished data that can modify the prognosis and diagnosis of liver metastases, as well as contribute to patients having new therapeutic options with improved survival.

Key words: liver metastasis; proteomics; mass spectrometry; prognostic biomarker; therapeutic resistance.

LISTA DE FIGURAS

Figura 1	Fluxograma representando as etapas desenvolvidas no estudo.....	25
Figura 2	Fluxograma PRISMA dos estudos de expressão gênica selecionados	27
Figura 3	Fluxograma representando as etapas de seleção das amostras FFPE rastreadas na instituição	34
Figura 4	PCA e <i>heatmap</i> das amostras de RNA- <i>seq</i> investigadas na meta-análise	45
Figura 5	Perfil global das 526 proteínas selecionadas após comparação entre metástases hepáticas inicialmente irressecáveis e tecidos adjacentes não-neoplásicos	50
Figura 6	<i>Heatmap</i> representando a atividade de sinalização entre metástases hepáticas irressecáveis e tecido não-neoplásico adjacente.....	54
Figura 7	Identificação de populações de células imunes nas amostras de metástase hepática e tecido não-neoplásico adjacente.....	56
Figura 8	Representação em <i>boxplot</i> da análise para confirmação do perfil de expressão das 14 proteínas selecionadas em cinco estudos independentes	57
Figura 9	Representação de amostras de metástases hepáticas, tumores primários do cólon e células mononucleares do sangue periférico obtidas por <i>single-cell</i> RNA- <i>seq</i>	58
Figura 10	Representação dos dados de <i>single-cell</i> RNA- <i>seq</i> de amostras de metástases hepáticas para a investigação de populações imunes e dos alvos selecionados anteriormente.....	59
Figura 11	Confirmação da expressão dos alvos selecionados em conjuntos independentes de amostras.....	61

Figura 12	Atividade das vias de sinalização dos estudos GSE3964, GSE104645, GSE53127 e GSE62322	65
Figura 13	Identificação de 32 proteínas diferencialmente abundantes após comparação entre os grupos de pacientes que apresentaram PD em tempo igual ou inferior a 6 meses (metástases resistentes) e pacientes que apresentaram PD após esse período (metástases sensíveis).....	66
Figura 14	<i>Heatmap</i> representando a atividade de sinalização entre metástases hepáticas irresssecáveis resistentes e sensíveis à terapia sistêmica	69
Figura 15	Perfil imune obtido com as proteínas selecionadas após comparação entre metástases hepáticas irresssecáveis resistentes e sensíveis à terapia sistêmica.....	71
Figura 16	Comparação de proteínas diferencialmente abundantes entre metástases irresssecáveis resistentes e sensíveis ao tratamento com 63 metástases colorretais com e sem recidiva após terapia obtidas no estudo GSE81423	72
Figura 17	Sobrevida global de acordo com a expressão dos alvos CDH2, CXCL14, SERPINA1 e LGR5	73

LISTA DE TABELAS

Tabela 1	Dados de expressão gênica de 20 estudos classificados de acordo com a plataforma utilizada e dados clínicos disponíveis.....	40
Tabela 2	Resumo das características clínicas e histopatológicas dos 15 pacientes com metástases hepáticas irresssecáveis do câncer de cólon incluídos nesse trabalho	41
Tabela 3	Representação dos 40 principais genes diferencialmente expressos entre metástases hepáticas e CCR _{M1}	46
Tabela 4	Enriquecimento para vias de sinalização com os 20 principais genes diferencialmente expressos com expressão aumenta identificados na comparação entre metástases hepáticas e CCR _{M1} utilizando a base de dados do KEGG	48
Tabela 5	Proteínas diferencialmente abundantes selecionadas após a comparação entre metástases inicialmente irresssecáveis do fígado e tecido adjacente.	51
Tabela 6	Enriquecimento para vias de sinalização a partir das proteínas identificadas como diferencialmente abundantes na comparação entre metástases hepáticas e tecido adjacente não-neoplásico utilizando a base de dados do KEGG.....	53
Tabela 7	Perfil de ativação e inibição de vias de sinalização considerando os dados de proteômica obtidos de metástases hepáticas inicialmente irresssecáveis e tecidos não-neoplásicos pareados.....	55
Tabela 8	Genes diferencialmente expressos entre metástases hepáticas do câncer de cólon resistentes e sensíveis ao tratamento	63
Tabela 9	Enriquecimento dos genes para vias de sinalização utilizando a base de dados do KEGG	65

Tabela 10	Identificação das 32 proteínas diferencialmente abundantes após comparação entre pacientes que apresentaram PD em tempo igual ou inferior a 6 meses (metástases hepáticas resistentes) e pacientes que apresentaram PD após período de 6 meses (metástases hepáticas sensíveis).....	67
Tabela 11	Enriquecimento para vias biológicas considerando as 32 proteínas diferencialmente abundantes entre metástases sensíveis e resistentes ao tratamento sistêmico.....	68
Tabela 12	Perfil de ativação e inibição de vias em metástases sensíveis e resistentes ao tratamento.....	70

LISTA DE SIGLAS E ABREVIATURAS

FDA	<i>Food and Drug Administration</i>
5-FU	5-Fluorouracil
AJCC	<i>American Joint Committee on Cancer</i>
CAP	Capecitabina
CAPIRI	Regime quimioterápico baseado em capecitabina e irinotecano
CAPOX	Regime quimioterápico baseado em capecitabina e oxaliplatina
CCD	Câncer de cólon direito
CCE	Câncer de cólon esquerdo
CCR	Câncer colorretal
CEP	Comitê de Ética em Pesquisa
CIMP	Fenótipo metilador das ilhas CpG (do inglês <i>CpG island methylator phenotype</i>)
CIMP-H	Níveis elevados do fenótipo metilador das ilhas CpG (do inglês <i>CpG island methylator phenotype-high</i>)
CIMP-L	Níveis intermediários do fenótipo metilador das ilhas CpG (do inglês <i>CpG island methylator phenotype-low</i>)
CIN	Instabilidade cromossômica (do inglês <i>chromosomal instability</i>)
CIN-H	Alta instabilidade cromossômica (do inglês <i>chromosomal instability-high</i>)
CMS	Consenso de subtipos moleculares (do inglês <i>consensus molecular subtypes</i>)
CPTAC	<i>Clinical Proteomic Tumor Analysis Consortium</i>
CR	Câncer retal
CRIS	Subtipos intrínsecos do câncer colorretal (do inglês <i>colorectal cancer intrinsic subtypes</i>)
CTCs	Células tumorais circulantes
ECM	Matriz extracelular (do inglês <i>extracellular matrix</i>)
EGA	<i>European Genome-Phenome Archive</i>
EGFR	Receptor do fator de crescimento epidérmico
EMT	Transição epitélio-mesênquima (do inglês <i>epithelial-mesenchymal transition</i>)

ESI	<i>Electrospray Ionization</i>
ESMO	<i>European Society for Medical Oncology</i>
FC	<i>Fold-change</i>
FDR	<i>False Discovery Rate</i>
FFPE	Amostras fixadas em formalina e embebidas em parafina (do inglês <i>formalin-fixed and paraffin-embedded</i>)
FGFR	Receptor do fator de crescimento de fibroblastos
FOLFIRI	Regime quimioterápico baseado em fluorouracil, leucovorina e irinotecano
FOLFIRINOX	Regime quimioterápico baseado em fluorouracil, leucovorina, oxaliplatina e irinotecano
FOLFOX	Regime quimioterápico baseado em fluorouracil, leucovorina e oxaliplatina
FOLFOXIRI	Regime quimioterápico baseado em fluorouracil, leucovorina, oxaliplatina e irinotecano
GTE_x	<i>Genotype-Tissue Expression Portal</i>
H&E	Hematoxilina-Eosina
IGFR	Receptor do fator de crescimento semelhante à insulina
IRI	Irinotecano
LC-MS	Cromatografia líquida acoplada à espectrometria de massas (do inglês <i>liquid chromatography-mass spectrometry</i>)
LV	Leucovorina
M0	Tumor primário do cólon não-metastático
M1	Tumor primário do cólon metastático
mAbs	Anticorpos monoclonais (do inglês <i>monoclonal antibodies</i>)
MALDI	Ionização por dessorção a laser assistida por matriz (do inglês <i>matrix-assisted laser desorption/ionization</i>)
mCCR	Câncer colorretal metastático
MET	Receptor do fator de crescimento de hepatócito
MMR	Sistema de reparo de mal pareamento do DNA (do inglês <i>Mismatch Repair</i>)
MS	Espectrometria de massas (do inglês <i>mass spectrometry</i>)
MSI	Instabilidade de microssatélites (do inglês <i>microsatellity instability</i>)

MSI-H	Alta instabilidade de microssatélites (do inglês <i>microsatellity instability-high</i>)
MSS	Estabilidade de microssatélites (do inglês <i>microsatellity stable</i>)
GEO	<i>Gene Expression Omnibus</i>
OX	Oxaliplatina
PCA	Análise de componentes principais (do inglês <i>principal component analysis</i>)
PC1	Componente principal 1 (do inglês <i>principal component 1</i>)
PD	Doença progressiva
PDGFR	Receptor do fator de crescimento derivado de plaquetas
PIGF	Fator de crescimento placentário
PR	Resposta parcial à quimioterapia
PRISMA	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analysis</i>
RECIST	<i>Response Evaluation Criteria in Solid Tumors</i>
RNA-seq	Sequenciamento de RNA (do inglês <i>RNA-sequencing</i>)
RTK	Receptores tirosina-quinases (do inglês <i>receptor tyrosine kinase</i>)
SD	Doença estável
TANs	Neutrófilos associados ao tumor (do inglês <i>tumor-associated neutrophils</i>)
TCGA	<i>The Cancer Genome Atlas</i>
Tis	Carcinoma <i>in situ</i>
TGF-β	Fator de transformação do crescimento beta (do inglês <i>transforming growth factor-β</i>)
TMB-H	Alta carga mutacional (do inglês <i>tumor mutational burden-high</i>)
TME	Microambiente tumoral (do inglês <i>tumor microenvironment</i>)
TMT	<i>Tandem Mass Tag</i>
TNF-α	Fator de necrose tumoral alfa (do inglês <i>tumor necrosis factor-α</i>)
TNM	Tamanho e/ou extensão do tumor primário, acometimento de linfonodos regionais e metástases à distância
TOF	<i>Time of Flight</i>
Tregs	Linfócitos T reguladores
UMI	<i>Unique Molecular Identifier</i>
VEGF	Fator de crescimento endotelial vascular
VEGFR	Receptor do fator de crescimento endotelial vascular

SUMÁRIO

1	INTRODUÇÃO.....	1
1.1	CÂNCER COLORRETAL.....	1
1.2	PROCESSO CARCINOGENICO	2
1.3	ESTADIAMENTO	4
1.4	CARACTERÍSTICAS ANATÔMICAS E MOLECULARES DO CCR.....	6
1.5	mCCR E METÁSTASES HEPÁTICAS	9
1.6	TRATAMENTO.....	12
1.7	RESISTÊNCIA TERAPÊUTICA.....	18
1.8	PROTEÔMICA DE METÁSTASES HEPÁTICAS	20
2	OBJETIVO.....	24
2.1	OBJETIVO GERAL	24
2.2	OBJETIVOS ESPECÍFICOS	24
3	METODOLOGIA.....	25
3.1	DESENHO DO ESTUDO	25
3.2	ANÁLISE TRANSCRIPTÔMICA E PROTEÔMICA <i>IN SILICO</i> DE AMOSTRAS DE CÂNCER DE CÓLON E METÁSTASES HEPÁTICAS	25
3.2.1	Meta-análise de dados públicos de expressão gênica.....	25
3.2.2	Casuística para integração de dados de RNA-seq.....	27
3.2.3	Casuística para integração de dados de <i>microarray</i>	28
3.2.4	Casuística para análise proteômica	29
3.2.5	Casuística para análise de <i>single-cell</i> RNA-seq.....	30
3.2.6	Processamento e análise dos dados	30
3.3	PROTEÔMICA DE METÁSTASES HEPÁTICAS INICIALMENTE IRRESSECÁVEIS DO CÂNCER DE CÓLON	32
3.3.1	Seleção das amostras.....	33
3.3.2	Casuística	34
3.3.3	Preparação das amostras para análise proteômica.....	35
3.3.4	Extração e digestão de proteínas	35
3.3.5	Análise do perfil proteômico por LC/MS	37
3.3.6	Parâmetros de processamento e busca em banco de dados públicos	37

4	RESULTADOS	39
4.1	CARACTERÍSTICAS GERAIS DOS DADOS PÚBLICOS DE EXPRESSÃO GÊNICA.....	39
4.2	DADOS CLÍNICO-PATOLÓGICOS DOS CASOS INCLUÍDOS NA ANÁLISE PROTEÔMICA.....	41
4.3	PERFIL TRANSCRIPTÔMICO E PROTEÔMICO GLOBAL DAS METÁSTASES 44	
4.4	PERFIL GLOBAL E PROTEÍNAS DIFERENCIALMENTE ABUNDANTES ENTRE METÁSTASES IRRESSECÁVEIS E TECIDOS ADJACENTES	49
4.5	ATIVIDADE DA SINALIZAÇÃO CELULAR	54
4.6	PERFIL IMUNOLÓGICO DAS AMOSTRAS INVESTIGADAS	55
4.7	COMPARAÇÃO ENTRE TRANSCRITOS E PROTEÍNAS E CONFIRMAÇÃO COM ESTUDOS INDEPENDENTES DA LITERATURA	56
4.8	PERFIL TRANSCRIPTÔMICO E PROTEÔMICO GLOBAL DAS METÁSTASES RESISTENTES E SENSÍVEIS À TERAPIA.....	62
4.9	ATIVIDADE DA SINALIZAÇÃO CELULAR E PERFIL IMUNE DE METÁSTASES	69
4.10	PERFIL IMUNOLÓGICO DAS AMOSTRAS INVESTIGADAS	70
4.11	COMPARAÇÃO DOS RESULTADOS EM CONJUNTOS INDEPENDENTES DE METÁSTASES HEPÁTICAS	71
4.12	IMPACTO DO PERFIL DE EXPRESSÃO DOS ALVOS NA SOBREVIVÊNCIA DOS PACIENTES	72
5	DISCUSSÃO.....	74
6	CONCLUSÃO.....	81
7	REFERÊNCIAS.....	83
8	ANEXOS.....	98
	Anexo 1 – Carta de aprovação do Comitê de Ética em Pesquisa (CEP)	98

1 INTRODUÇÃO

1.1 CÂNCER COLORRETAL

O câncer colorretal (CCR) é uma neoplasia de origem epitelial que abrange tumores que se desenvolvem no cólon (subdividido em ceco, cólon ascendente, flexura hepática, cólon transversal, flexura esplênica, cólon descendente e cólon sigmoide), transição retossigmoide e reto. É um dos cânceres mais incidentes e prevalentes na população mundial, sendo o segundo mais comum entre as mulheres e o terceiro mais comum entre os homens (GLOBOCAN 2021). Segundo as estimativas globais para 2020, cerca de 2 milhões de novos casos e 935 mil mortes foram previstas, correspondendo a aproximadamente um em cada 10 casos de câncer, tanto por ocorrência quanto por óbito (GLOBOCAN 2021). Desconsiderando o agrupamento dos sítios anatômicos cólon, reto e ânus, o câncer de cólon permanece responsável por 6,3% dos novos casos e aproximadamente 577 mil mortes (5,8%), excluindo-se os casos de câncer de pele não melanoma, assim apresentando maiores taxas de incidência e mortalidade quando comparado ao câncer de reto (GLOBOCAN 2021).

A porcentagem de incidência do câncer de cólon tende a variar dependendo da região, com maiores taxas nas regiões europeias, Austrália/Nova Zelândia e na América do Norte, e menores taxas na África e na região centro-sul da Ásia (GLOBOCAN 2021). Essa variação pode ser relacionada ao desenvolvimento socioeconômico dos países, em que taxas mais altas são observadas nos países mais desenvolvidos, e regiões em desenvolvimento apresentam uma tendência no aumento de novos casos acompanhando o crescimento do Índice de Desenvolvimento Humano, como notado em muitos países da Europa Oriental, Ásia e América do Sul (Dekker et al. 2019; GLOBOCAN 2021). Para o Brasil, estimou-se cerca de 41 mil novos casos de CCR em 2020, sendo 20.540 em homens e 20.470 em mulheres, representando 9,1% e 9,2% do total de todas as neoplasias para cada sexo, respectivamente (Ministério da Saúde 2020).

Aproximadamente 70% dos casos de CCR são esporádicos, ou seja, sem história familiar ou aparente predisposição genética (Kuipers et al. 2015). Além de estarem associados a uma maior prevalência em indivíduos do sexo masculino e com o aumento da idade, sendo mais frequente na população com idade igual ou superior a 50 anos,

dispõem de uma etiologia multifatorial que envolve fatores modificáveis como obesidade, tabagismo, consumo excessivo de álcool e de carnes vermelhas e processadas (Brenner et al. 2014; Simon 2016; Thanikachalam e Khan 2019). Entretanto, fatores de risco não modificáveis também estão associados à doença, como história pessoal de pólipos, doença inflamatória intestinal, presença de *diabetes mellitus* tipo 2 e resistência à insulina (Dekker et al. 2019; Simon 2016; Thanikachalam e Khan 2019). Os demais casos são decorrentes de agrupamentos familiares positivos e de mutações em genes associados a síndromes hereditárias características que possivelmente interagem com fatores ambientais e influenciam na evolução da doença (Kuipers et al. 2015).

De forma inversa, sugere-se que a prática de atividade física, suplementação de cálcio e mudanças dietéticas, como consumo de peixes, grãos inteiros, fibras e laticínios, refletem na redução do risco de se desenvolver CCR (Dekker et al. 2019). Apesar dos dados alarmantes, o CCR é um dos cânceres com maiores possibilidades de ser evitado e tratado se detectado precocemente através dos métodos de rastreamento (Jeon et al. 2018), considerando que o estadiamento ao diagnóstico é o preditor mais importante de sobrevida com taxas relativas de 5 anos variando entre 90% a 14% (Siegel et al. 2020). As diretrizes de triagem normalmente se baseiam na idade do indivíduo (entre 45-50 e 75 anos), histórico de doença inflamatória intestinal e histórico pessoal e/ou familiar relacionado à presença de pólipos ou tumores colorretal (Jeon et al. 2018; Thanikachalam e Khan 2019). Contudo, observa-se um aumento de novos casos durante as últimas décadas na população com idade inferior a 50 anos (Siegel et al. 2020) e, somado a isso, a ausência de sinais e sintomas e a falta de adesão de muitos pacientes junto às medidas de rastreamento precoce prejudicam o diagnóstico de estádios iniciais.

Em vista disso, o CCR configura-se como um problema de saúde pública devido à sua agressividade e particularidade e, apesar de existir pouco entendimento sobre sua etiologia, sua carcinogênese ocorre pelo acúmulo de alterações genéticas que geram instabilidade cromossômica e desequilíbrio de funções celulares importantes (Penney et al. 2020).

1.2 PROCESSO CARCINOGENICO

Durante o processo carcinogênico, células normais são transformadas em células tumorais através do acúmulo de mutações e da seleção natural, em que esse processo

permite a adaptação dessas células a novos ambientes, adquirindo fenótipos malignos relacionados, por exemplo, ao desenvolvimento de metástases e resistência ao tratamento (Niida et al. 2021). Em 1976, Nowell foi o pioneiro em propor que o câncer se desenvolve através desse sistema evolucionário e descobertas subsequentes levaram ao primeiro modelo do processo carcinogênico colorretal descrito por Fearon e Vogelstein (1990). Os autores caracterizaram a sequência adenoma-carcinoma, na qual ocorre a transformação de células epiteliais normais em uma lesão benigna (adenoma) e, posteriormente, em um tumor maligno através do acúmulo de sucessivas mutações fundadoras em genes *drivers* como *APC*, *KRAS*, *TP53*, *SMAD4* e *DCC* (Fearon e Vogelstein 1990; Niida et al. 2021; Sarvestani et al. 2020).

Com o advento de novas técnicas genômicas, hoje sabe-se que a heterogeneidade intratumoral está relacionada à ramificação de um clone ancestral comum em subclones que adquirem mutações progressivas ao longo do tempo (Niida et al. 2021). As mutações fundadoras, presentes nas células-tronco neoplásicas da base das criptas colônicas, se acumulam durante as fases iniciais e estão presentes em todas as regiões da massa tumoral (Dekker et al. 2019; Niida et al. 2021). Por sua vez, mutações progressivas estão presentes em algumas dessas regiões, formando perfis mutacionais distintos (Niida et al. 2021).

Diante disso, a forma esporádica do CCR normalmente se inicia a partir de uma cripta aberrante que evolui para uma lesão precursora e que pode eventualmente progredir para um tumor entre um período de 10 a 15 anos (Dekker et al. 2019). Atualmente é conhecido que, além da tradicional via adenoma-carcinoma, as lesões precursoras, também chamadas de pólipos, podem se desenvolver através de uma via alternativa serrilhada. Os pólipos adenomatosos provindos da via tradicional representam cerca de 60-85% dos CCRs, e os 15-40% remanescentes ocorrem, em sua maioria, através da via da neoplasia serrilhada com a sequência pólipo serrilhado-adenocarcinoma (Dekker et al. 2019; Nakanishi et al. 2019).

Somado a isso, os tumores colorretais e suas respectivas lesões precursoras se desenvolvem a partir de vias moleculares distintas. A via da instabilidade cromossômica (CIN, do inglês *chromosomal instability*), geralmente associada aos adenomas tradicionais, é observada em cerca de 70% dos diagnósticos e caracterizada pelo acúmulo sequencial de múltiplas mutações durante as divisões celulares (Simon 2016). O fenótipo de CIN é iniciado por uma mutação inativadora no gene supressor tumoral *APC*, afetando a segregação cromossômica durante a fase mitótica, seguida por ativação de *RAS* e

subsequente perda de função do *TP53* e *SMAD4* (Dekker et al. 2019; Simon 2016; Nakanishi et al. 2019).

Em contrapartida, a mutação ativadora em *BRAF* é considerada a principal alteração genética que promove as lesões originadas pela via serrilhada, em especial a alteração no éxon 15 do gene onde ocorre uma substituição do aminoácido valina por um ácido glutâmico no códon 600 (*BRAF*^{V600E}) e afeta o domínio quinase da proteína (Nakanishi et al. 2019). Outra mutação oncogênica comum nessa via é em *KRAS* (tipicamente nos códons 12 e 13), em que ambas as ativações resultam na estimulação constitutiva da via MAPK e consequentemente desregula a proliferação, diferenciação e sobrevivência celular (Nakanishi et al. 2019). Além desse evento, as lesões serrilhadas também são caracterizadas pela instabilidade epigenética devido ao fenótipo metilador de ilhas CpG (CIMP, do inglês *CpG island methylator phenotype*), que ocasiona a hipermetilação da região promotora de diversos genes supressores tumorais como *MLH1*, *BMP3* e *NDRG4* (Dekker et al. 2019; Simon 2016; Nakanishi et al. 2019).

Através de estudos em maior escala, o fenótipo relacionado à instabilidade de microssatélites (MSI, do inglês *microsatellity instability*) também foi identificado em tumores que apresentam mutações em genes do sistema de reparo de mal pareamento do DNA (MMR, do inglês *mismatch repair*), podendo ocorrer nas vias de ambas as lesões (Dekker et al. 2019). A MSI está associada tanto com mutações germinativas nos genes MMR quanto com mutações somáticas relacionadas à metilação da região promotora do gene *MLH1* (Simon 2016). O silenciamento de *MLH1* é significativamente associado com mutações em *BRAF* e indica tumores com altos níveis de hipermetilação (CIMP-H) e com alta MSI (MSI-H), enquanto que tumores com alterações em *KRAS* apresentam o gene *MLH1* intacto, com níveis intermediários do fenótipo CIMP (CIMP-L) e normalmente com estabilidade de microssatélites (MSS, do inglês *microsatellity stable*) (McInnes et al. 2017; Nakanishi et al. 2019).

1.3 ESTADIAMENTO

O sistema de classificação TNM (tamanho e/ou extensão do tumor primário, acometimento de linfonodos regionais e metástases à distância), estabelecido pelo *American Joint Committee on Cancer* (AJCC), é o parâmetro prognóstico mais robusto na estratificação de pacientes de acordo com as características histopatológicas dos tumores (Weiser 2018). Na 8ª edição do protocolo, divulgado em outubro de 2016

(Chicago, IL, Estados Unidos) e implementado globalmente em janeiro de 2018, houve algumas mudanças significativas a respeito do carcinoma *in situ*, referido como adenocarcinoma intramucoso, considerando passível de ocorrência de metástase apenas tumores colorretais que invadem a submucosa (Weiser 2018). Entretanto, a principal diferença entre AJCC-7ª e AJCC-8ª é que o sistema de estadiamento CCR foi revisado para incluir um novo estágio envolvendo metástase peritoneal (denominado estágio IVC). Outra importante adição no protocolo é a ênfase na medicina personalizada, com foco em painéis moleculares para pesquisa de proficiência dos genes MMR ou de MSI como fatores prognósticos e o status de *BRAF*, *KRAS* e *NRAS* como um preditor de eficácia terapêutica (Weiser 2018).

De forma geral, a classificação patológica permanece a mesma, descrevendo a profundidade de invasão do tumor primário através da parede do cólon (T), número de linfonodos regionais acometidos (N) e presença de metástase à distância (M). A partir dessas observações, os tumores podem ser classificados em 4 diferentes estágios com prognóstico e decisões terapêuticas distintas (Sagaert et al. 2018; Weiser 2018). No estágio I os pacientes apresentam uma taxa de sobrevivência de 5 anos de 94%, passando para 82% no estágio II em que os tumores podem se desenvolver através das camadas mais externas do órgão e invadir tecidos adjacentes. O estadiamento III, com uma taxa de 67%, já indica comprometimento dos linfonodos regionais e pacientes estadiado IV apresentam metástases à distância, com uma taxa relativa de 5 anos de apenas 11% (Sagaert et al. 2018).

Em ambas as classificações, AJCC-7ª e 8ª, o componente N é considerado um fator prognóstico mais importante do que o componente T, considerando que T1N1a é estadiado como IIIA, enquanto T4bN0 é estadiado como IIC (Zhang et al. 2020). No entanto, estudos mostraram que os componentes T e N têm importância comparável, visto que os tumores T4N0 tiveram resultados significativamente piores do que os tumores T1-2N1-2a, independentemente do número de linfonodos examinados (Chan et al. 2018; Chu et al. 2016). Isso sugere que a discriminação e a precisão das avaliações prognósticas usando a classificação TNM da AJCC permanecem controversas, especialmente para os estágios II e III do CCR, levando a pedidos de uma classificação TNM modificada (mTNM).

1.4 CARACTERÍSTICAS ANATÔMICAS E MOLECULARES DO CCR

Outro fator conhecido por influenciar no prognóstico dos pacientes é a localização das neoplasias. Apesar da denominação conjunta, o câncer de cólon difere do retal em questões etiológicas, genéticas e moleculares, além do manejo terapêutico dependendo do estadiamento e da resposta clínica (Gao et al. 2017; Mukund et al. 2020). De acordo com a posição anatômica dos tumores, o CCR pode ser classificado em câncer de cólon direito (CCD ou proximal) e câncer de cólon esquerdo (CCE ou distal), separados pela flexura esplênica, e câncer retal (CR), sendo este último referente a lesões localizadas nos 12 cm da borda anal (Gao et al. 2017). Embora ainda controverso se os três tipos podem ser considerados uma única entidade ou três distintas, há evidências na literatura de que o CR compartilha características comuns com o CCE, e ambos diferem geneticamente do CCD (TCGA 2012).

O CCE geralmente apresenta adenocarcinomas típicos com histologia característica da via carcinogênica tradicional e, consequentemente, com alta CIN (CIN-H, do inglês *chromosomal instability-high*) (Baran et al. 2018). Por apresentarem essa morfologia, o aspecto polipoide das lesões facilita sua detecção precoce, o que favorece a sobrevida do paciente (Baran et al. 2018; TCGA 2012). Além disso, é mais frequente em homens e dependente da atividade de receptores tirosina-quinases (RTK, do inglês *receptor tyrosine kinase*), incluindo a sinalização do receptor do fator de crescimento epidérmico (EGFR) (Ghidini et al. 2018). Por sua vez, o CCD apresenta maior ocorrência em pacientes com idade mais avançada e do sexo feminino (Jiang et al. 2020), além de normalmente se apresentarem como adenocarcinomas mucinosos com MSI-H (TCGA 2012). Esses tumores apresentam pior prognóstico, difícil detecção em estádios iniciais por sua morfologia plana e, consequentemente, são maiores e mais agressivos, geralmente pouco diferenciados (Baran et al. 2018). Para este último, associou-se um pior desfecho clínico e ineficácia às terapias anti-EGFR devido a alterações moleculares, incluindo mutações no gene *BRAF* (Baran et al. 2018; Gao et al. 2017).

A análise dos perfis moleculares dos CCRs tem sido responsável pela geração de uma grande quantidade de dados, permitindo que as mutações sejam melhor caracterizadas nesses tumores (Zheng et al. 2020). Tumores com CIN apresentam um perfil não-hipermutado e, de acordo com o estudo realizado pelo *The Cancer Genome Atlas* (TCGA), além das mutações já identificadas que ocorrem na sequência adenoma-carcinoma, variantes nos genes *PIK3CA*, *FBXW7*, *TCF7L2*, *CTNNB1*, *SMAD2* e

FAM123B também são frequentemente encontradas (TCGA 2012). Por outro lado, o perfil hipermutado é encontrado em tumores com o fenótipo de CIMP e MSI, com mutações frequentes nos genes *ACVR2A*, *APC*, *TGFBR2*, *MSH3*, *MSH6*, *SLC9A9* e *TCF7L2* em conjunto com *BRAF*^{V600E} (TCGA 2012).

Vários genes mutados ou com a função de seus produtos proteicos alterada estão envolvidos na desregulação de vias de sinalização importantes na iniciação e desenvolvimento do CCR (Koveitypour et al. 2019). Por exemplo, a inativação de *APC* leva a uma ativação excessiva da via WNT e, em acordo com a alta frequência de mutações nesse gene, essa via foi encontrada alterada em 93% das amostras tumorais de CCR analisadas pelo consórcio TCGA (Buikhuisen et al. 2020; TCGA 2012). Ainda, as vias PI3K e TGF- β estão frequentemente alteradas em estádios avançados por alterações genômicas nos genes *PIK3CA* e *PTEN*, e *SMAD2* e *SMAD4*, respectivamente (Buikhuisen et al. 2020). Outras vias como EGFR/MAPK e Notch estão associadas à regulação de vários processos biológicos, incluindo proliferação celular, diferenciação, angiogênese e apoptose (Kennedy et al. 2020; Pranteda et al. 2020).

Em decorrência dos avanços no rastreamento molecular em grande escala e exaustivas análises computacionais, Guinney et al. (2015) consolidaram os resultados de diferentes grupos de pesquisa e estabeleceram o consenso de subtipos moleculares (CMS, do inglês *consensus molecular subtypes*), definindo um sistema de estratificação baseado nos diferentes níveis de expressão gênica dos CCRs entre os estágios I a IV (Buikhuisen et al. 2020). Os autores descreveram quatro subtipos (CMS1 a CMS4) que, além de diferirem entre si pelas características biológicas e moleculares dos tumores, são associados ao desfecho clínico dos pacientes a partir de diferentes taxas de sobrevida global e livre de progressão (Guinney et al. 2015).

O subtipo CMS1 é frequentemente diagnosticado em mulheres com CCD, caracterizando assim tumores com uma alta carga mutacional (TMB-H, do inglês *tumor mutational burden-high*) por apresentarem MSI-H e CIMP-H, maior quantidade de neoantígenos e alta infiltração de células imunes (Buikhuisen et al. 2020; Guinney et al. 2015; Ouyang et al. 2021). Tumores com os subtipos CMS2 e CMS3 compartilham uma alta expressão de uma assinatura epitelial, porém aqueles CMS2 são principalmente encontrados no lado esquerdo do cólon e também são denominados como subtipo canônico por apresentarem altos níveis de CIN e ativação das vias de sinalização WNT e MYC (Buikhuisen et al. 2020; Guinney et al. 2015). Em contrapartida, o CMS3 também é descrito como subtipo metabólico pela desregulação de diversas vias relacionadas ao

metabolismo celular e acompanhado por mutações no gene *KRAS* (Buikhuisen et al. 2020; Guinney et al. 2015). Por fim, CMS4 apresenta altos níveis de expressão de genes relacionados a um fenótipo mesenquimal, com ativação do processo denominado transição epitélio-mesênquima (EMT, do inglês *epithelial-mesenchymal transition*) e das vias de sinalização TGF- β e angiogênese, com abundante componente estromal e infiltração de linfócitos e monócitos (Buikhuisen et al. 2020; Guinney et al. 2015). Embora a lateralidade e o status de mutações em *RAS* ou *RAF* sejam fatores que auxiliam na escolha de tratamentos sistêmicos, a classificação CMS está sendo explorada em ensaios clínicos como um marcador prognóstico ou preditivo, sendo que os CCDs são mais frequentemente MSI-ímmunes (CMS1) e metabólicos (CMS3), enquanto os CCEs normalmente estão associados aos subtipos canônico (CMS2) e mesenquimal (CMS4) (Dekker et al. 2019). Além disso, foram observadas diferenças na distribuição dos quatro subtipos dentro de diferentes estágios do sistema TNM, destacando o aumento de pacientes com subtipo mesenquimal (CMS4) nos estágios mais avançados da doença (Coebergh van den Braak et al. 2020).

Além das classificações propostas, Isella et al. (2017) incluíram uma estratificação secundária baseada no CMS explorando subtipos intrínsecos do CCR (CRIS, do inglês *colorectal cancer intrinsic subtypes*) a partir de modelos de xenoenxertos derivados de pacientes. A finalidade, tanto em desenvolver o CRIS quanto de usar uma estratégia *in vivo*, era de estabelecer subtipos relacionados ao transcriptoma específico de células neoplásicas do CCR, excluindo a influência dos componentes do microambiente tumoral (TME, do inglês *tumor microenvironment*) que podem gerar diferentes classificações CMS mesmo em amostras obtidas do mesmo paciente (Isella et al. 2017). A partir do CRIS, cinco assinaturas intrínsecas foram estabelecidas (CRIS-A a CRIS-E), em que estas se sobrepõem significativamente às descritas previamente no CMS (Buikhuisen et al. 2020). CRIS-A consiste em tumores com *KRAS* mutado ou fenótipo MSI e mutação em *BRAF*, enquanto CRIS-B representa amostras pouco diferenciadas com ativação das vias TGF- β e EMT. O subtipo CRIS-C engloba tumores com CIN-H e com *KRAS* selvagem, amplificação de *MYC* e ativação da via de sinalização EGFR. CRIS-D apresenta ativação constitutiva das vias de sinalização WNT, receptor do fator de crescimento de fibroblastos (FGFR) e do receptor do fator de crescimento semelhante à insulina (IGFR). Por fim, CRIS-E também é caracterizado pela alta atividade de WNT, com maior frequência de tumores com mutações em *TP53* (Buikhuisen et al. 2020; Isella et al. 2017).

Apesar desses esforços e avanços nas análises multiômicas, as classificações propostas ainda não foram implementadas na prática clínica com uma influência direta nas decisões terapêuticas, seja pela complexidade de aprovar e implementar algoritmos, custos da operação ou pelo impacto na rotina da equipe médica (Valenzuela et al. 2021). Somado a isso, um grande desafio no tratamento desses tumores é a natureza heterogênea da doença que dificulta, principalmente, a adoção de terapias precisas. Essa heterogeneidade provém do envolvimento de diversas vias de sinalização e alterações moleculares exclusivas, acarretando em comportamentos e respostas distintas mesmo quando alocados no mesmo estadió patológico ou subtipo molecular. Essas características compreendem diferenças em tumores do mesmo tipo histológico e que podem ser observadas a partir de subpopulações diferentes no tumor primário ou entre o tumor primário e suas lesões metastáticas, respectivamente (Sagaert et al. 2018).

1.5 mCCR E METÁSTASES HEPÁTICAS

A doença metastática é caracterizada pela identificação de tumores secundários em órgãos distantes ao CCR (mCCR). Cerca de 20% dos indivíduos já apresentam metástases ao diagnóstico (estádio IV, metástase sincrônica) e até 40% dos pacientes com doença localizada (estádio I-III) tratados previamente com abordagem cirúrgica desenvolvem recidiva metastática dentro de um período de cinco anos (metástase metacrônica) (Biller e Schrag 2021; Piawah e Venook 2019; Poturnajova et al. 2021). Aproximadamente 90% dos pacientes com estadió IV apresentam metástases no fígado, pulmão, cérebro ou peritônio (Poturnajova et al. 2021). Considerando pacientes não-metastáticos, a probabilidade de ocorrência desse evento é inferior a 10% para o estágio I (extensão do tumor primário até a camada muscular própria, linfonodo negativo), entre 10% a 20% para o estágio II (se estende através da muscular própria ou em estruturas adjacentes, linfonodo negativo) e até 50% para o estágio III (linfonodo positivo) (Biller e Schrag 2021).

O fígado é o local mais comum para o desenvolvimento de metástases devido ao sistema venoso portal que fornece abundante suprimento sanguíneo entre os dois órgãos (Tsilimigras et al. 2021). Entre 30% a 50% dos indivíduos com CCR desenvolverão lesões hepáticas durante o curso da doença, sendo que metade desses pacientes apresentam doença sincrônica (Tsilimigras et al. 2021). Brevemente, esse processo se inicia com um subconjunto de células do tumor primário que passa por alterações morfológicas através

do programa EMT, perdendo características epiteliais pela modificação da expressão de moléculas de adesão e adquirindo características que permitem a migração através da matriz extracelular (ECM, do inglês *extracellular matrix*) e invasão de tecidos adjacentes (Zhou et al. 2022). Após o intravasamento de vasos sanguíneos, as células tumorais circulantes (CTCs) atingem o fígado e extravasam de sua vasculatura, colonizando o órgão secundário (Tsilimigras et al. 2021; Zhou et al. 2022). Durante essa fase, as células disseminadas precisam reverter o fenótipo mesenquimal adquirido para recuperarem seu potencial proliferativo, sendo um processo crítico que pode resultar em morte celular, ativação de um estado de dormência (quiescência celular) ou a efetiva proliferação, estabelecendo micrometástases que se expandem até formarem tumores clinicamente detectáveis (macrometástases) (Poturnajova et al. 2021; Tsilimigras et al. 2021; Zhou et al. 2022).

O complexo estabelecimento e desenvolvimento das metástases hepáticas envolve diversos mecanismos moleculares, além de estar diretamente relacionado ao TME e sua interação com o ambiente externo ao tumor (Zhou et al. 2022). Na via hematogênica, as CTCs são auxiliadas por plaquetas e neutrófilos para evadirem da imunovigilância das células NK, e já na microvasculatura hepática, essas células são capazes de iniciar uma cascata pró-inflamatória que aciona a secreção de quimiocinas pelas células residentes altamente especializadas do fígado e permite a adesão das CTCs às células endoteliais para posterior extravasamento e colonização (Poturnajova et al. 2021). A prévia modulação do órgão secundário também é fundamental para estabelecer um nicho pré-metastático que favoreça a sobrevivência e o crescimento dessas células, incluindo o recrutamento de células progenitoras hematopoiéticas que expressam receptores do fator de crescimento endotelial vascular (VEGFR) do tipo 1 (VEGFR-1) e que auxiliam a neoangiogênese, além de células inflamatórias e imunes para ativação de processos que visam a remodelação tecidual e imunossupressão (Poturnajova et al. 2021; Zhou et al. 2022). Além disso, algumas vias de sinalização possuem um papel crucial no processo metastático, como a WNT, EGFR, TGF- β , Notch e HGF/c-MET (Koveitypour et al. 2019; Poturnajova et al. 2021; Zhou et al. 2022).

Apesar de toda a caracterização genômica em grande escala realizada pelo TCGA e a associação desses dados com o rastreamento proteômico e análises proteogenômicas geradas pelo *Clinical Proteomic Tumor Analysis Consortium* (CPTAC), um importante projeto desenvolvido pelo *National Cancer Institute* em 2011, o conhecimento sobre a base genética das metástases hepáticas provenientes do câncer de cólon ainda não foi

completamente elucidado (Ma et al. 2019). A comparação entre tumores primários e metastáticos aponta que, embora existam características semelhantes entre os tecidos, as metástases desenvolvem assinaturas genéticas específicas inerentes ao próprio programa do órgão secundário que caracteriza seu perfil de expressão (Hartung et al. 2017; Menck et al. 2022). Paralelamente, a heterogeneidade intertumoral entre o CCR e suas respectivas metástases pode ser um fator limitante quando considerado a escolha da terapia-alvo com base no status mutacional dos genes *KRAS*, *NRAS* e *BRAF* do tumor primário. Mesmo com dados contraditórios na literatura, em que autores identificaram um alto grau de similaridade nas alterações genômicas desses genes *drivers* (Liu et al. 2020a; Sayagués et al. 2016; Vakiani et al. 2012), Del Carmen et al. (2018) avaliaram essa questão e detectaram apenas 56% de concordância entre metástases hepáticas e linfonodais comparadas à doença primária, o que reforça a necessidade de se considerar o tecido metastático durante as análises quando se trata de medicina de precisão.

Estudos realizados nos últimos anos obtiveram sucesso em identificar alterações no transcriptoma entre tumores primários e metastáticos, destacando genes com papel central em redes de interação gênica ou proteica e em assinaturas associadas à presença e infiltração de células imunes (Liu et al. 2021; Zhang et al. 2019; Zhang et al. 2021). Além das assinaturas expressas exclusivamente nos tecidos metastáticos, recentemente foi demonstrado que o subconjunto de células do CCR que adquire propriedades invasivas e migratórias perde seu programa genético específico do tecido primário antes da colonização do órgão secundário, podendo expressar genes característicos do tecido-alvo, mas em níveis inferiores aos encontrados nas metástases (Teng et al. 2020). Assim, as metástases hepáticas passam por uma reprogramação dessas assinaturas, com um aumento de expressão de genes específicos do fígado e perda daqueles que caracterizam o epitélio intestinal (Teng et al. 2020).

Somado a isso, outras evidências sugerem um enriquecimento significativo de vias metabólicas e de processos relacionados à angiogênese, inflamação e imunossupressão nas metástases (Liu et al. 2020a; Kamal et al. 2019; Kim et al. 2019; Zhou et al. 2021), mostrando uma modulação específica que contribui para o estabelecimento e sobrevivência do tumor no fígado. Apesar dos avanços das ferramentas de bioinformática terem proporcionado uma maior integração dos dados, possibilitando novas descobertas acerca do CCR e mCCR, a investigação e disponibilização pública das informações de metástases hepáticas ainda permanecem limitadas. Considerando todas as modificações transcricionais e moleculares necessárias durante o complexo processo de

inicialização e estabelecimento dos novos tumores secundários, a compreensão de assinaturas e mecanismos intrínsecos desses tumores representa um grande potencial no avanço da medicina personalizada.

1.6 TRATAMENTO

Além do diagnóstico clínico e anatomopatológico que auxilia no estadiamento do paciente de acordo com o sistema TNM, a avaliação da expressão das proteínas dos genes MMR e a pesquisa de mutações em *KRAS*, *NRAS* e *BRAF* também são empregadas para estratificar pacientes com o objetivo de atingir melhores resultados terapêuticos (Biller e Schrag 2021; Li et al. 2020; Punt et al. 2017). A cirurgia e a quimioterapia são as duas principais opções de tratamento para o câncer de cólon, dependendo do estágio e da localização do tumor no momento do diagnóstico, bem como das características individuais dos pacientes (Patel et al. 2019).

O tratamento ideal para o CCR visa a remoção completa do tumor e de suas metástases, quando aplicável. Assim, a ressecção cirúrgica é a única ou principal abordagem empregada nos estadiamentos I-III, podendo ser associada à quimioterapia adjuvante para pacientes com estágio II com alto risco de recidiva e para todos diagnosticados com estágio III (Xie et al. 2020). Em contrapartida, o mCCR é definido como ressecável quando o tumor primário e todas as metástases são passíveis de remoção cirúrgica completa, que normalmente ocorre quando a doença secundária está limitada a alguns focos, tipicamente no fígado ou pulmão (Biller e Schrag 2021). Para aqueles pacientes com lesões irresssecáveis ou que não são elegíveis à cirurgia, a radioterapia e quimioterapia são as principais estratégias utilizadas com o objetivo de reduzir o tamanho do tumor e impedir seu crescimento e disseminação (Xie et al. 2020).

Atualmente o regime quimioterápico pode ser ofertado em agente único, principalmente à base de fluoropirimidinas utilizando o fluorouracil (5-FU) por via intravenosa ou o seu pró-fármaco capecitabina (CAP) por via oral, ou em combinação com um ou vários medicamentos, incluindo oxaliplatina (OX) e irinotecano (IRI) (Biller e Schrag 2021; Xie et al. 2020). O 5-FU foi o primeiro medicamento com atividade significativa contra o CCR, sendo posteriormente associado ao ácido fólico (leucovorina/LV) por este apresentar efeito modulador e sinérgico ao quimioterápico (Piawah e Venook 2019). Embora estudos tenham mostrado que a terapia de primeira linha com agente único não é inferior aos regimes combinados quando avaliado a

sobrevida global dos pacientes, as combinações denominadas FOLFOX (5-FU+LV+OX), FOLFIRI (5-FU+LV+IRI), CAPOX (CAP+OX) e CAPIRI (CAP+IRI) são as principais abordagens utilizadas no tratamento sistêmico, reservando a quimioterapia de agente único para pacientes intolerantes ou que apresentaram baixo desempenho clínico (Biller e Schrag 2021; Xie et al. 2020).

Para pacientes com metástases hepáticas, o tratamento cirúrgico é o único que oferece chance de cura e sobrevida de 5 e 10 anos ao redor de 40% e 25%, respectivamente, comparando-se com uma taxa de sobrevida de cerca de 5% para pacientes tratados de forma paliativa (Adam e Kitano 2019; Chow e Chok 2019; Vera et al. 2020). Contudo, apesar das indicações cirúrgicas terem aumentado drasticamente nas últimas décadas devido ao progresso nas abordagens e surgimento de novas técnicas, apenas 20% dos pacientes são considerados elegíveis para a cirurgia no momento do diagnóstico, que depende principalmente de fatores que envolvem a quantidade, localização e capacidade de ressecção completa das lesões, presença de metástases extra-hepáticas não-ressecáveis, volume remanescente do fígado e do estado geral do paciente (Adam et al. 2020).

Uma alternativa a essa limitação seria alcançar a ressecabilidade da doença hepática através do tratamento sistêmico. A base de dados internacional LiverMetSurvey (<https://livermetsurvey-arcad.org/>), que conta com mais de 70 países participantes, avaliou 4.000 pacientes com doença inicialmente irressecável e que foram operados após receberem quimioterapia neoadjuvante, mostrando resultados promissores com uma taxa de sobrevida de 5 anos de 32% (Adam e Kitano 2019). Outros estudos com a mesma proposta reportaram uma taxa de resposta entre 50% a 64% após a quimioterapia sistêmica, com conversão favorável de 20-40% dos pacientes (Lam et al. 2012; Martin et al. 2020). No entanto, apesar de todos os esforços visando maiores chances de cura e aumento da sobrevida, mais da metade dos pacientes que passaram pela ressecção da lesão metastática desenvolvem recidiva, principalmente da doença intra-hepática, com uma concentração maior de casos nos primeiros 5 anos de seguimento clínico (Biller e Schrag 2021; Martin et al. 2020; Tsilimigras et al. 2018; Zarour et al. 2017).

O mCCR permanece incurável para a maior parte dos pacientes. Sem tratamento, a sobrevida é inferior a um ano, porém pesquisas realizadas nas últimas décadas apontam uma extensão desse período para quase 20 meses quando ofertada quimioterapia sistêmica paliativa (Chow e Chok 2019; Piawah e Venook 2019; Zhou et al. 2022). As fluoropirimidinas (5-FU e CAP) têm sido utilizadas há mais de 50 anos como tratamento

de primeira linha para CCR estágio IV, presentes nas combinações FOLFOX/CAPOX ou FOLFIRI/CAPIRI (Biller e Schrag 2021). Os regimes podem ser intercambiáveis quando as taxas de respostas não são satisfatórias ou pela ocorrência de toxicidade severa, que geralmente engloba neuropatia sensorial com o uso de OX e mais diarreia e alopecia com IRI (Biller e Schrag 2021). Pacientes com estágio III tratados com FOLFOX adjuvante e que desenvolveram metástases metacrônicas podem responder à reintrodução da combinação quimioterápica, porém quando essa ocorre dentro de 12 meses após a exposição ou quando se observa neuropatia persistente, o FOLFIRI acaba sendo a melhor estratégia (Biller e Schrag 2021). A combinação tripla denominada FOLFOXIRI ou FOLFIRINOX (5-FU+LV+OX+IRI) também pode ser considerada para pacientes com alta carga tumoral e doença progressiva a fim de alcançar resultados mais promissores, embora essa terapia intensa também leve ao aumento da toxicidade, incluindo neutropenia, fadiga, diarreia e neuropatia (Biller e Schrag 2021; Chow e Chok 2019).

Mais recentemente houve um progresso no tratamento dos pacientes com mCCR devido à introdução da terapia-alvo. Essa modalidade engloba agentes direcionados às características biológicas de cada tumor, podendo atuar em antígenos de superfície, fatores de crescimento, receptores ou em vias de transdução de sinal, bloqueando assim sinais celulares que favorecem a sobrevivência e proliferação das células tumorais. Além disso, o TME e seus constituintes também podem ser ativados por essas drogas a fim de fortalecer a vigilância e ataque imunológico (Lee et al. 2018; Piawah e Venook 2019). Os anticorpos monoclonais (mAbs, do inglês, *monoclonal antibodies*) são os principais agentes nessas terapias direcionadas, compondo um grupo de moléculas pequenas com peso molecular inferior a 900 Da que permite o reconhecimento tanto de alvos intra quanto extracelulares (Lee et al. 2018). Várias vias que controlam a iniciação, progressão e migração do CCR, como WNT, Notch, Hedgehog e TGF- β /SMAD, bem como aquelas capazes de ativar cascatas de sinalização, incluindo PI3K/AKT e MAPK (RAS/RAF/MEK/ERK), contêm locais ideais para terapia direcionada (Krishnamurthy et al. 2018; Xie et al. 2020). Dada a complexa sinalização *downstream* e as dificuldades em inibir interações biológicas específicas, nem todas as vias existentes relacionadas ao CCR podem ser interferidas com sucesso, e os dados atuais cobrem apenas algumas nas quais os agentes direcionados identificados experimentalmente podem ser comprovados como eficientes em estudos clínicos (Tiwari et al. 2018; Xie et al. 2020).

A seleção de pacientes para diferentes modalidades de tratamento utilizando um agente biológico é feita a partir da detecção de mutações ativadoras nos genes da família

RAS e *BRAF*. Esses oncogenes codificam proteínas envolvidas na via clássica de sinalização MAPK, resultando em sua ativação constitutiva que consequentemente favorece a proliferação e sobrevivência das células tumorais. As mutações são encontradas em cerca de 35-45% (*KRAS*), 4% (*NRAS*) e 5-9% (*BRAF*) dos mCCR (Gbolahan e O'Neil, 2019; Shi et al. 2020). O EGFR é altamente expresso no CCR (25-77%) e, uma vez ativado, é responsável pelo acionamento de diversas vias de sinalização celular como MAPK, PI3K/AKT e JAK/STAT3 (Xie et al. 2020). Assim, pacientes *RAS* e *RAF* selvagem podem ser tratados com mAbs anti-EGFR como parte da primeira linha de tratamento quimioterápico, podendo seguir com o uso do agente anti-VEGF (fator de crescimento endotelial vascular) (Adam e Kitano 2019). Entretanto, apenas este último é uma opção para aqueles que apresentam mutações ou status mutacional desconhecido, visto que esse fator permite a ativação constitutiva de MAPK e outras vias correlatas independentemente da ligação de fatores de crescimento aos RTK (Adam e Kitano 2019). Ao adicionar essas terapias direcionadas aos regimes ofertados para pacientes com mCCR, maiores taxas de resposta (>50%) e expectativa de vida (cerca de 30 meses) foram alcançadas (Adam e Kitano 2019).

O primeiro agente direcionado para CCR aprovado pela *Food and Drug Administration* (FDA) foi o cetuximabe em 2004, seguido por bevacizumabe no mesmo ano, e a partir de então medicamentos emergentes aprovados para o CCR foram sucessivamente introduzidos no mercado (Xie et al. 2020). Atualmente, o cetuximabe e panitumumabe são os dois agentes biológicos anti-EGFR aprovados pela FDA para uso no tratamento de primeira linha do mCCR (Wu et al. 2020). O ensaio ASPECCT fase III verificou que ambos apresentam efeitos farmacológicos similares, não havendo inferioridade entre os dois medicamentos (Price et al. 2014; Xie et al. 2020). Apesar disso, o uso desses agentes anti-EGFR não é atrativo para tratamentos de segunda linha, visto que geralmente não há resultados significativos no tempo de sobrevida livre de progressão ou global (Xie et al. 2020). Assim, as terapias anti-VEGF são as melhores opções para esses pacientes.

O bevacizumabe é um mAb humanizado direcionado ao VEGF-A, mostrando melhora na sobrevida tanto global quanto livre de progressão nos casos de mCCR independentemente do status mutacional de *KRAS* nos pacientes (Wu et al. 2020). Em termos de eficácia, não há diferenças nos benefícios clínicos alcançados entre sua associação com FOLFOX ou FOLFIRI (Xie et al. 2020). No entanto, uma melhor resposta foi observada quando o agente anti-VEGF foi combinado ao regime FOLFOXIRI,

mostrando que a intensificação do tratamento de primeira linha traz melhores resultados, porém acompanhados com o aumento de eventos adversos hematológicos e gastrointestinais (Cremolini et al 2020; Xie et al. 2020). Sua administração é relevante tanto na primeira quanto na segunda linha de tratamento, possibilitando uma melhor resposta clínica com o aumento de taxas de sobrevida em pacientes que progrediram aos regimes iniciais (Wu et al. 2020). Paralelamente, o agente também pode ser utilizado no período de manutenção de pacientes que alcançaram doença estável após o tratamento inicial, melhorando drasticamente o tempo e a qualidade de vida desses indivíduos (Xie et al. 2020). Além do bevacizumabe, outra opção de terapia antiangiogênica para a segunda linha de tratamento é o medicamento aflibercepte, que age como um antagonista bloqueando a ligação do VEGF-A, VEGF-B e fator de crescimento placentário (PIGF) aos receptores extracelulares VEGFR-1 e -2 (Xie et al. 2020). Apesar de seus limitados benefícios como agente único, sua combinação com FOLFIRI em pacientes que receberam previamente OX com ou sem adição de bevacizumabe mostrou melhoras significativas na sobrevida global e livre de progressão, além de maiores taxas de resposta (Xie et al. 2020).

Como supracitado, embora a investigação da lateralidade do tumor primário não esteja incluída no consenso publicado pela *European Society for Medical Oncology*, as diretrizes do *National Comprehensive Cancer Network* recomendam o uso de substâncias anti-EGFR para o tratamento de mCCR provindo do CCE com *RAS* selvagem (Baek 2017; Messersmith 2019). A identificação da mutação *BRAF*^{V600E} também é importante, pois torna os tumores mais agressivos e esses não respondem de forma satisfatória à terapia sistêmica, podendo-se adotar outras estratégias como a inclusão de um inibidor de *BRAF* (anti-*BRAF*) (Baek 2017). A eficácia desses inibidores no melanoma, um tipo de câncer com frequências mais altas de *BRAF* mutado do que o CCR, incentivou o desenvolvendo de uma abordagem semelhante para o mCCR, sendo necessário associar os anti-*BRAFs* a inibidores *upstream* da via MAPK para evitar a reativação por *feedback* dos RTK. Assim, resultados promissores foram alcançados quando se associou o vemurafenibe ou encorafenibe (anti-*BRAFs*) às terapias já utilizadas (Xie et al. 2020).

Outras abordagens são empregadas para pacientes com mCCR refratários aos regimes quimioterápicos propostos. O regorafenibe é um inibidor que tem como alvo múltiplas tirosina-quinases angiogênicas, incluindo o VEGFR, FGFR e o receptor do fator de crescimento derivado de plaquetas (PDGFR), bem como as proteínas RET, KIT e *BRAF* (Grothey et al. 2013). Os estudos clínicos CORRECT e CONCUR validaram sua

eficácia no tempo médio de sobrevida dos pacientes não-responsivos, publicando resultados significativos quando comparado seu uso ao placebo (Grothey et al. 2013; Li et al. 2015). Ainda, o TAS-102 é uma monoterapia administrada por via oral que combina a trifluridina com o cloridrato de tipiracila e representa uma nova opção para pacientes com mCCR que já passaram por, pelo menos, duas linhas de tratamento (Cicero et al. 2020). Além dos seus efeitos benéficos na sobrevida e no controle dos sintomas da doença, com vantagens que independem do status mutacional de *RAS* e *RAF*, o TAS-102 também pode ser ofertado para pacientes com idade mais avançada, sendo bem tolerado por estes não desenvolverem efeitos adversos significativos que resultem na interrupção do tratamento (Cicero et al. 2020).

Um recente avanço na terapia dos mCCRs foi a aprovação pela FDA dos inibidores do *checkpoint* imunológico, uma terapia direcionada que visa aumentar a vigilância imunológica e supressão contra as células tumorais, bloqueando sua evasão da detecção pelas células T (Xie et al. 2020). Atualmente existem dois mAbs humanizados que têm como alvo PD-1 e que são aprovados pela FDA, o pembrolizumabe e nivolumabe (Le et al. 2015; Xie et al. 2020). A via da PD-1 regula negativamente a resposta imune citotóxica às células neoplásicas. Assim, os tumores que possuem regulação positiva de PD-1 e PD-L1 escapam do sistema imune do hospedeiro. Devido às deficiências no MMR, tumores MSI-H têm quantidades aumentadas de mutações, neoantígenos e linfócitos infiltrantes tumorais, assim apresentando maiores níveis de expressão de PD-1, PD-L1 e CTLA-4 (Kalyan et al. 2018). Apesar de resultados promissores com imunoterapia, apenas 5% dos pacientes com mCCR são MSI-H e nem todos com esse perfil apresentam a mesma resposta à terapia oferecida. Em pacientes candidatos a imunoterapia, estudos recentes mostram que essas diferenças nos perfis de respostas podem ser atribuídas à heterogeneidade do TME (Pedrosa et al. 2019).

As opções de tratamento expandidas, associando a combinação de medicamentos quimioterápicos com agentes biológicos, melhoraram a sobrevida global mediana de aproximadamente um ano com o uso de 5-FU para cerca de dois a três anos (Cremolini et al. 2015). Entretanto, embora os avanços científicos tenham contribuído fortemente para a seleção de tumores candidatos a imunoterápicos a partir de características genômicas, ainda permanece um desafio desvendar as propriedades que levam tumores com perfis semelhantes a desenvolver resistência, destacando a importância de novas estratégias investigativas nessa área.

1.7 RESISTÊNCIA TERAPÊUTICA

Apesar dos avanços no diagnóstico e tratamento do mCCR, as taxas de mortalidade pela doença ainda permanecem mundialmente altas. Essa estatística pode ser associada à apresentação de um perfil de resistência à terapia que resulta na inefetividade do tratamento e ocorre em 90% dos pacientes com doença avançada (Taieb et al. 2019; Wang et al. 2019). Esse fenômeno pode ser mediado por diversos mecanismos celulares e moleculares, incluindo mutações secundárias e ativação de vias compensatórias, que acabam favorecendo a sobrevivência das células tumorais por estas se tornarem tolerantes ao tratamento farmacológico (Taieb et al. 2019; Wang et al. 2019). A resistência pode ser categorizada em dois tipos dependendo do momento em que se desenvolve: intrínseca ou primária, que já está presente antes da oferta dos regimes quimioterápicos, e adquirida ou secundária, que é induzida após a administração e se manifesta com uma gradual redução da eficácia do tratamento (Wang et al. 2019; Yeoh et al 2021). Coletivamente, esses fatores produzem um cenário no qual os pacientes tratados com as terapias aprovadas e disponíveis atualmente obtêm benefícios de sobrevida global e sem progressão na escala de semanas a meses, mas raramente por um tempo superior.

A desregulação de vias de sinalização desempenha um papel crítico dentro dos inúmeros mecanismos moleculares que contribuem para a aquisição desse perfil de resistência (Yeoh et al 2021). As vias MAPK, PI3K/AKT, WNT e Notch já foram destacadas com papéis relevantes nesse aspecto e, além disso, evidências recentes também sugerem o envolvimento de um *crosstalk* complexo e controlado entre diferentes vias de transdução de sinal (He et al. 2018; Yeoh et al 2021). Paralelamente, a presença de subpopulações celulares que retêm diferentes características genéticas define a heterogeneidade intratumoral e pode ser um fator limitante em termos de eficácia terapêutica (Wang et al. 2019). Assim, pacientes que apresentam subclones resistentes intrinsecamente podem responder à terapia em um primeiro momento pela apresentação de uma população majoritária sensível, mas após a interrupção do tratamento desenvolvem recidiva pela retomada da proliferação dessas células remanescentes (Wang et al. 2019).

Conforme descrito anteriormente, a genotipagem dos genes *KRAS*, *NRAS* e *BRAF* tem um grande impacto no manejo clínico e na seleção do tratamento de primeira linha dos pacientes com mCCR, visto que o status mutacional desses genes se tornou um preditor validado de resposta à terapia anti-EGFR (Blank-Landeshammer et al. 2019;

Franczak et al. 2020). Entretanto, apesar da verificação da mutação em *KRAS* ter se tornado cotidiano na prática clínica, menos de 30% dos pacientes que possuem o gene em status selvagem respondem de forma satisfatória à terapia (Blank-Landeshammer et al. 2019; Bray et al. 2019; Wang et al. 2019). Sugere-se que alterações em outras proteínas participantes da cascata de sinalização EGFR, como PIK3CA e PTEN, podem contribuir para a redução da sensibilidade ao cetuximabe através da via PI3K/AKT, além da amplificação em *MET* que foi detectada em pacientes que inicialmente responderam ao tratamento, mas eventualmente adquiriram resistência (Bray et al. 20129; Wang et al. 2019; Yeoh et al 2021). Além disso, a ativação de vias compensatórias foi demonstrada com o uso do bevacizumabe, por exemplo, o qual inibindo VEGF desencadeou a sinalização a partir do receptor do fator de crescimento de hepatócito (*MET*), FGFR, IGFR do tipo 1 ou PDGFR (Blank-Landeshammer et al. 2019).

Em aproximadamente 10% dos pacientes com mCCR que abrigam mutações em *BRAF*^{V600E} é observado pior prognóstico e ativação constitutiva da via de sinalização de MAPK (Paton et al. 2020). Normalmente, BRAF e outras quinases RAF são ativadas pelos membros da família RAS (*KRAS*, *NRAS* ou *HRAS*) em resposta aos sinais dos RTK. Uma vez ativado, BRAF e outras quinases RAF fosforilam e ativam as quinases MEK (*MEK1* e *MEK2*), que por sua vez fosforilam e ativam as quinases ERK (*ERK1* e *ERK2*) (Paton et al. 2020). Embora em tumores como melanoma, portadores de mutações *BRAF*^{V600E}, tenham sido observadas taxas de resposta maiores que 50% aos anti-BRAF, a eficácia tem sido decepcionante no mCCR *BRAF*^{V600E}, com taxas de resposta de apenas 5% (Caputo et al. 2019). Estudos pré-clínicos sugeriram que a resistência ao anti-BRAF no mCCR é impulsionada pela reativação da sinalização MAPK através da perda de sinais de *feedback* negativo, levando à ativação do RAS por meio dos RTK (Corcoran et al. 2012; Prahallad et al. 2012). Com base nesses achados, ensaios clínicos recentes avaliaram combinações de anti-BRAF com anti-EGFR ou -MEK, ou a combinação tripla dos inibidores de BRAF, EGFR e MEK em pacientes com CCR *BRAF*^{V600E}, com o objetivo de produzir uma inibição mais eficaz e duradoura de MAPK (Corcoran et al. 2018; Tabernero et al. 2021). Apesar dessas combinações terem levado a taxas de resposta modestamente melhoradas nos CCRs, essas respostas não apresentam grande duração e o benefício clínico é limitado pelo rápido surgimento de resistência adquirida (Corcoran et al. 2018).

Considerando a extensa caracterização ao longo dos últimos anos por meio do rastreamento de transcritos e mutações somáticas associadas à progressão e desfecho

clínico dos pacientes com mCCR, ainda se observam inconsistências entre os dados obtidos e o fenótipo tumoral. Além disso, muitas vezes a tomada de decisão para tratamento da doença metastática ainda é baseada nas características do tumor primário, desconsiderando aquelas específicas das metástases que podem influenciar diretamente as respostas terapêuticas, como a composição do TME, assinaturas moleculares características e a própria heterogeneidade intertumoral entre o mCCR e essas lesões. Em vista disso, a análise direcionada para esses tumores pode representar um grande avanço científico na área.

1.8 PROTEÔMICA DE METÁSTASES HEPÁTICAS

A natureza complexa do câncer torna um desafio se basear apenas nas informações genômicas e transcriptômicas considerando o manejo clínico dos pacientes, principalmente porque esses aspectos podem não se correlacionar fielmente à abundância de proteínas, que passam por modificações traducionais e pós-traducionais que ocasionam alterações no produto proteico esperado em comparação aos níveis de mRNA (Blank-Landeshammer et al. 2019; Rodriguez et al. 2021; Zhang et al. 2014). Considerando que essas moléculas são alvos da maioria das intervenções terapêuticas atuais e em desenvolvimento (Rodriguez et al. 2021), a investigação proteômica das células tumorais traz novas percepções sobre o comportamento molecular envolvendo assinaturas relacionadas à resistência terapêutica e dinâmica de vias de sinalização.

A proteômica é o estudo do conjunto de proteínas de uma amostra biológica visando, entre outras coisas, a identificação e quantificação de suas abundâncias de acordo com diferentes condições (Lindhorst e Hummon 2020). Nas últimas décadas, têm sido inestimável a importância da técnica analítica de espectrometria de massas (MS, do inglês *mass spectrometry*) no avanço de pesquisas envolvendo a identificação do perfil global de proteínas e disponibilização dos achados em bancos de dados robustos. Além do projeto CPTAC, que acelerou a compreensão da base molecular do câncer, outra importante iniciativa foi o desenvolvimento do projeto *The Human Protein Atlas*, iniciado em 2000 e que, atualmente, conta com um extenso acervo de proteínas expressas em células normais, tecidos e tumores (<https://www.proteinatlas.org>).

Na análise por MS, a detecção e quantificação de proteínas é baseada na relação massa/carga de íons gerados em fase gasosa e pode ser realizada por duas principais

estratégias: *top-down*, que avalia proteínas intactas, e *bottom-up* ou *shotgun*, que visa a avaliação de peptídeos obtidos pela digestão do material proteico com o uso de uma enzima proteolítica, inferindo a presença de proteínas pela comparação das sequências peptídicas identificadas com a sequência do proteoma de referência (O'Neill 2019; Rozanova et al. 2021). Somado a isso, a diferença entre as estratégias que envolvem a digestão proteica é que a *shotgun* requer uma prévia separação físico-química da mistura complexa de peptídeos através da cromatografia líquida acoplada à MS (LC-MS, do inglês *liquid chromatography-mass spectrometry*), produzindo os espectros proporcionalmente à eluição do analito (Matthiesen e Bunkenborg 2013; O'Neill 2019). Brevemente, o espectrômetro de massas é um equipamento que integra uma fonte de ionização, analisador de massas, detector de íons e sistema de aquisição de dados, podendo-se utilizar diferentes técnicas e abordagens para cada etapa, como a fonte de íons por ESI (do inglês *electrospray ionization*) ou pela ionização por dessorção a laser assistida por matriz e os analisadores quadrupolo e TOF (do inglês *time of flight*), bem empregados e descritos na literatura (Calderón-Celis et al. 2018).

Nos últimos anos, grande parte da literatura se concentra na investigação proteômica do CCR a partir de amostras frescas (congeladas), fixadas em formalina e embebidas em parafina (FFPE, do inglês *formalin-fixed and paraffin-embedded*) e biópsias líquidas (Lindhorst e Hummon 2020). Rotineiramente, os hospitais armazenam a maioria dos tecidos biopsiados que são coletados para fins de diagnóstico e cortes de peças cirúrgicas na forma de FFPE, visto que essa técnica é uma escolha econômica que permite tanto o armazenamento quanto a preservação da integridade de uma grande quantidade de material à temperatura ambiente por anos ou décadas (Coscia et al 2020; Mitsa et al. 2022). Além disso, o arquivamento a longo prazo possibilita a análise retrospectiva desses espécimes, associando e integrando novos achados a uma infinidade de dados clínicos incluindo relatórios histológicos, tratamentos ofertados e respostas terapêuticas (Coscia et al 2020). Os recentes e contínuos avanços em técnicas e protocolos mostraram sucesso na aquisição de dados proteômicos de pequenas quantidades de amostras e destacaram a potencial utilização de amostras FFPE como uma alternativa à investigação de tecidos frescos congelados, apresentando propriedades qualitativas e quantitativas semelhantes frente à técnica de MS (Coscia et al 2020).

Diversos estudos já exploraram o perfil de proteínas abundantes das metástases hepáticas, estabelecendo alvos exclusivos ou diferenciais envolvidos em processos biológicos e vias de sinalização participantes do fenótipo tumoral. Considerando a base

molecular do tumor primário, enquanto Fahrner et al. (2021) reportaram um distinto proteoma que se estende à heterogeneidade intertumoral entre amostras pareadas de mCCR e metástases, Ku et al. (2019) identificaram perfis de abundância similares, com uma clara separação das amostras apenas quando considerado os alvos diferencialmente abundantes entre os dois tecidos (Wong et al. 2022). Apesar dessas inconsistências, outros autores adotaram a mesma abordagem e destacaram alterações na abundância de proteínas relacionadas ao metabolismo energético (Fahrner et al. 2021; Kim et al. 2019), à diversas vias associadas a ECM (Liu et al. 2020b; Naba et al. 2014; van Huizen et al. 2019; Voß et al. 2019) e ao sistema imune (Fahrner et al. 2021; Kim et al. 2019; Liu et al. 2020), assim contribuindo para o conhecimento de proteínas diferenciais nesses tumores em relação ao mCCR e correlacionando-as a processos regulatórios fundamentais e interações proteína-proteína (Wong et al. 2022). Somado a isso, mais recentemente um estudo demonstrou alterações em moléculas associadas à atividade e metabolismo de fármacos nas metástases hepáticas, com diminuição da abundância de enzimas metabolizadoras e desregulação de transportadores e receptores que levam a perturbações em vias relevantes no processo de desenvolvimento e progressão do câncer (Vasilogianni et al. 2022).

Considerando todo o cenário apresentado, a caracterização genômica e molecular permitiu um grande avanço na terapia de pacientes com doença metastática. Além disso, o avanço das ferramentas de bioinformática proporcionou uma maior integração dos dados, possibilitando novas descobertas acerca do CCR e mCCR. Apesar de todos os esforços, a maior parte dos estudos se concentra na investigação de metástases hepáticas ressecáveis, o que pode refletir parcialmente os complexos mecanismos envolvidos na configuração e no comportamento da doença irresssecável, presente em cerca de 80% dos pacientes. Paralelamente, além desses indivíduos possuírem limitadas opções terapêuticas devido ao desenvolvimento de resistência, o pouco conhecimento molecular da doença torna um desafio a seleção precisa de proteínas no tumor que modulam a dinâmica da atividade de vias e caracterizam a doença como uma entidade específica. Dessa maneira, essa perspectiva favorece a hipótese de que essas células tumorais desenvolvem mecanismos de regulação que modulam a dinâmica da sinalização oncogênica por rotas secundárias e/ou estabelecimento de novas interações entre proteínas, mantendo sua atividade e evitando morte celular. Assim, a investigação proteômica global das metástases no fígado tem grande potencial para desvendar mecanismos intrínsecos que estejam

associados a falhas terapêuticas e que possam impactar na terapia dos pacientes com aumento da sobrevida.

2 OBJETIVO

2.1 OBJETIVO GERAL

Investigar e identificar moduladores da sinalização celular no câncer de cólon e metástases hepáticas que estejam associados ao desenvolvimento e progressão da doença.

2.2 OBJETIVOS ESPECÍFICOS

- Rastrear o perfil de expressão de transcritos e abundância de proteínas em mCCR e metástases hepáticas a partir de dados *in silico*;
- Avaliar o perfil proteômico global por LC-MS de metástases hepáticas irresssecáveis;
- Investigar o status de ativação e inibição de vias de sinalização nos dados gerados pela análise proteômica e disponíveis na literatura;
- Investigar mecanismos de resistência associados às proteínas selecionadas.

3 METODOLOGIA

3.1 DESENHO DO ESTUDO

Considerando as diferentes abordagens utilizadas, tanto computacionais quanto para o rastreamento proteômico das amostras, a metodologia foi organizada em dois grandes módulos, apresentados na **Figura 1**.

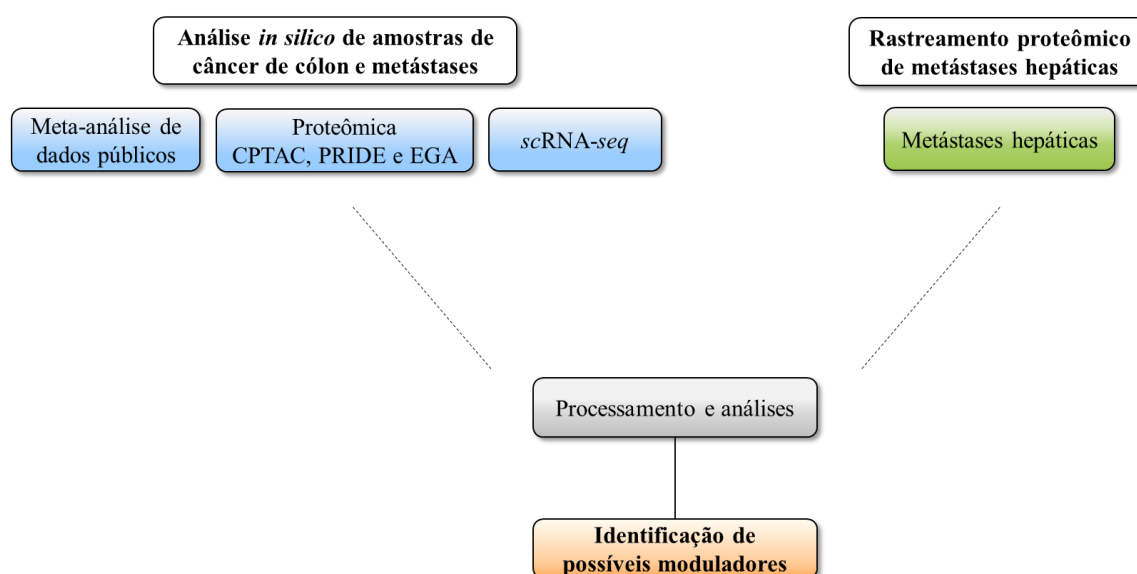


Figura 1 - Fluxograma representando as etapas desenvolvidas no estudo. A análise *in silico* envolveu o rastreamento de estudos públicos de transcriptômica e proteômica de amostras de câncer de cólon e metástases hepáticas. O rastreamento proteômico por espectrometria de massas foi feito em 30 biópsias FFPE pareadas do A.C.Camargo Cancer Center, sendo 15 de metástases hepáticas inicialmente irressecáveis e 15 de tecidos não-neoplásicos adjacentes. Esses dados foram analisados visando um melhor entendimento das características moleculares e sinalização desses tumores, permitindo assim a identificação de possíveis moduladores envolvidos tanto na biologia da doença quanto na resistência ao tratamento. CPTAC: *Clinical Proteomic Tumor Analysis Consortium*; PRIDE: *Proteomics Identifications Database*; EGA: *European Genome-Phenome Archive*; scRNA-seq: *single-cell RNA sequencing*; FFPE: amostras fixadas em formalina e embebidas em parafina.

3.2 ANÁLISE TRANSCRIPTÔMICA E PROTEÔMICA *IN SILICO* DE AMOSTRAS DE CÂNCER DE CÓLON E METÁSTASES HEPÁTICAS

3.2.1 Meta-análise de dados públicos de expressão gênica

As buscas por dados de transcriptoma foram realizadas nos repositórios públicos *Gene Expression Omnibus* (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) e *ArrayExpress*

(<https://www.ebi.ac.uk/arrayexpress/>) entre julho e agosto de 2021. Para ambos os casos, as buscas foram feitas utilizando os termos “colon”, “colorectal”, “metastasis”, “metastatic”, “liver”, “hepatic”, “expression profiling by array”, “expression profiling by high throughput sequencing” e suas respectivas combinações. Adicionalmente, utilizou-se da busca em artigos indexados na base de dados PubMed com os termos “colon cancer”, “colorectal cancer” e “Gene Expression Omnibus” para selecionar estudos utilizados na literatura e que não foram encontrados durante o rastreamento. Os resultados foram pré-filtrados de modo a incluir apenas estudos conduzidos com as técnicas de *microarray* ou sequenciamento de RNA (RNA-seq), com amostras de tumores primários esporádicos e metástases provenientes do CCR. Foram excluídos estudos duplicados ou cujos dados tenham sido obtidos a partir de linhagens celulares e modelos xenográficos. Por fim, a seleção dos estudos elegíveis incluiu aqueles com: a) amostras de câncer de cólon ou CCR que apresentavam informações sobre a localização anatômica das lesões; b) amostras de metástases hepáticas; c) dados clínicos informando, pelo menos, a avaliação do componente M do sistema TNM; d) avaliação de tumores metastáticos (M1); e) matrizes de contagem de *reads* disponíveis. A metodologia descrita, baseada no método PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analysis*) (Page et al. 2021), está representada na **Figura 2**.

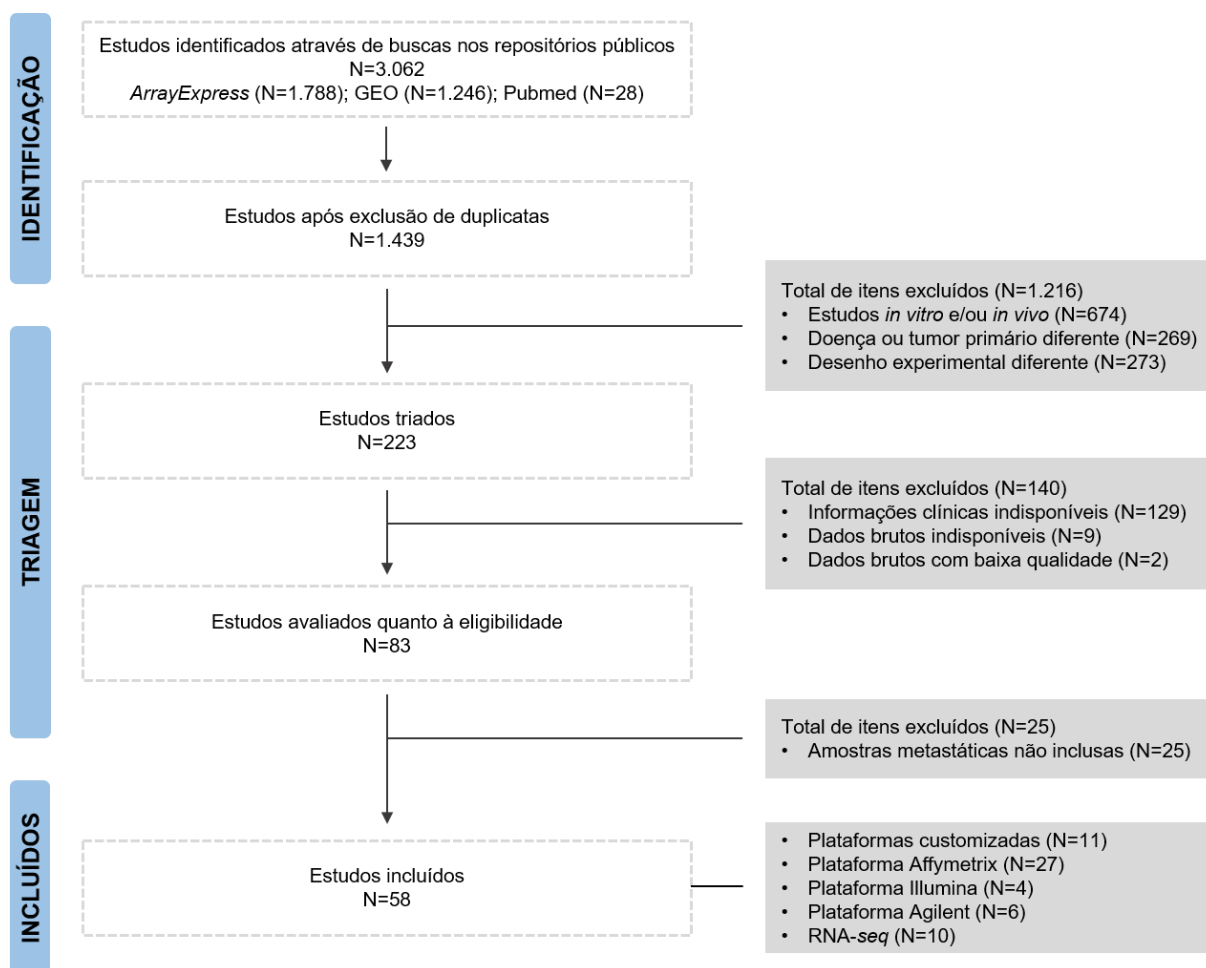


Figura 2 - Fluxograma PRISMA dos estudos de expressão gênica selecionados.

3.2.2 Casuística para integração de dados de RNA-seq

Para a análise de transcritos, 416 amostras de câncer de cólon (N=377) e tecido adjacente não neoplásico (N=39) geradas pelo consórcio internacional TCGA foram investigadas. A matriz de contagem de *reads* com HTSeq (versão *harmonized*) dessas amostras foi obtida pelo pacote TCGAbiolinks (v2.22.4) disponível no programa R/Bioconductor (Colaprico et al. 2016). A versão *harmonized* garante que os dados foram alinhados a última versão do genoma de referência disponível (GRCh38/hg38). Os dados clínicos, além de baixados pelo pacote mencionado, também foram obtidos do material suplementar do trabalho publicado por Liu et al. (2018) com informações de *follow-up* atualizadas. Foram mantidos apenas adenocarcinomas de cólon e excluídos aqueles localizados na junção reto-sigmóide. Em relação ao componente N do sistema TNM, 226 pacientes não apresentaram disseminação regional das células tumorais e 151 apresentaram. Por sua vez, avaliando a presença de metástases à distância (componente

M do sistema TNM), 279 pacientes eram não-metastáticos (M0) e 53 M1 – em 45 amostras essa categoria não pode ser avaliada (MX ou sem informação clínica).

Além dos dados gerados pelo TCGA, matrizes de contagem de *reads* de tecidos normais do cólon (N=779) e fígado (N=226) foram obtidas do Portal GTEx (*Genotype-Tissue Expression Portal*, <https://gtexportal.org/home/>) considerando a última versão disponibilizada (V8). Essa plataforma contém dados de transcrição gênica de indivíduos saudáveis que doaram os seus órgãos *post-mortem*, realizando a análise de RNA-seq com o protocolo Illumina TruSeq, pré-processando os dados com o alinhamento ao genoma de referência GRCh38 (hg38) e quantificando a contagem de *reads* com RNA-SeQC (v1.1.9) (GTEx Consortium 2013).

Por fim, também foram investigadas 58 amostras independentes de tumores primários do cólon (N=44) e metástases hepáticas (N=14) obtidas dos estudos de RNA-seq (N=7) eleitos elegíveis na etapa anterior de meta-análise e depositados no banco de dados GEO. Desses estudos, 3 foram realizados pela plataforma Illumina HiSeq 2000, 2 pela Illumina HiSeq 4000, e as plataformas Illumina NextSeq 500 e Illumina HiSeq 2000 foram utilizadas em apenas um estudo, respectivamente. As matrizes de contagem de *reads* e os dados clínicos referentes a cada estudo foram baixados manualmente do banco de dados GEO.

3.2.3 Casuística para integração de dados de *microarray*

Considerando a análise dos estudos de *microarray*, foram excluídos aqueles realizados em plataformas customizadas (N=11), com baixo número de amostras (N=2) e com arquivos de intensidade apresentando valores diferentes do padrão esperado (N=3). Dessa forma, 32 estudos de *microarray* foram baixados manualmente dos bancos de dados públicos sob os formatos “.txt” e “.CEL”. Considerando diferenças de qualidade que estão relacionadas à integridade do material armazenado, esses dados de *microarray* foram separados para integrar separadamente aqueles obtidos de amostras frescas (N=29) dos de amostras FFPE (N=3). Ainda, os estudos com amostras coletadas a fresco de metástases hepáticas de origem colorretal foram separadas daquelas oriundas de tecido colônico, que foram mantidas com os demais dados. Cada estudo foi corrigido considerando o *background* e as sondas foram anotadas para obter os identificadores únicos de genes a partir do banco de dados Entrez Gene (NCBI). Foram removidas as sondas que não anotaram para nenhum identificador Entrez, e foi obtido a média de

intensidades para aquelas que anotaram para o mesmo identificador. Dez estudos foram excluídos da análise por apresentarem um número inferior de identificadores. Entrez comparado com os demais. Assim, 3 tabelas foram geradas considerando a integração dos valores de intensidade dos estudos selecionados, sendo uma de amostras coletadas a fresco englobando metástases do câncer de cólon (N=16), outra contando com metástases do CCR (N=5) e a última contabilizando as amostras FFPE (N=3) – 2 estudos ficaram duplicados entre as duas primeiras tabelas. O primeiro conjunto de dados apresenta 1.699 amostras, sendo 1.445 de tumores primários, 50 de tecidos adjacentes não-neoplásicos e 204 metástases. A segunda tabela conta apenas com amostras de metástases hepáticas (N=141). Por fim, a tabela de amostras FFPE dispõe de 125 amostras de tumores primários e 63 de metástases.

3.2.4 Casuística para análise proteômica

Foram investigadas 173 amostras de câncer de cólon (N=73) e tecido adjacente não neoplásico (N=100) obtidas a partir do conjunto de dados proteômicos disponibilizados pelo consórcio CPTAC (Edwards et al. 2015, Ellis et al. 2013). Os autores realizaram a quantificação por LC-MS pelo método de marcação isobárica TMT (do inglês, *Tandem Mass Tag*), e as abundâncias relativas de 8.067 proteínas foram transformadas em $\log_2$ e recentralizadas para alcançar uma mediana comum de zero (Vasaikar et al. 2019). Todos os dados foram obtidos a partir do repositório CPTAC (<https://proteomics.cancer.gov/>). Da mesma forma que na transcriptômica, tumores classificados histologicamente como adenocarcinoma mucinoso e localizados na junção retossigmóide foram excluídos da análise. De acordo com a presença de metástases, 39 pacientes eram linfonodo negativo e 34 positivo em relação a disseminação regional de células tumorais, e, considerando a disseminação à distância, 38 amostras eram de pacientes M0 e 8 de M1 – em 27 amostras essa categoria não pôde ser avaliada (MX).

Com o objetivo de realizar comparações secundárias, também foram avaliados estudos de proteômica de metástases hepáticas do CCR disponibilizados em bancos de dados públicos. Foram pré-selecionados 5 estudos, dos quais apenas 2 prosseguiram devido aos critérios de inclusão/exclusão propostos anteriormente e disponibilidade de dados. Para a obtenção do estudo EGAD00010002237, disponível no repositório EGA (*European Genome-Phenome Archive*, <https://ega-archive.org/>), foi estabelecida uma colaboração com o pesquisador Dr. Oliver Schilling, do Institute of Surgical Pathology,

Medical Center University of Freiburg (Alemanha), que disponibilizou os arquivos não-processados (.mzML) para análise de metástases hepáticas ressecáveis (N=21) e tumores primários pareados (N=21). Por fim, arquivos não-processados (.mzML) de 14 metástases hepáticas ressecáveis também foram obtidos do estudo PXD008383, que está disponível no banco de dados PRIDE (*Proteomics Identifications Database*, www.ebi.ac.uk/pride/archive/).

3.2.5 Casuística para análise de *single-cell* RNA-seq

Os dados de *single-cell* RNA-seq foram obtidos do estudo GSE178318, que apresenta 15 amostras de seis pacientes, sendo 6 de metástases ressecáveis do fígado, 6 de câncer primário no cólon e 3 de sangue periférico coletado antes da cirurgia. Cada trio de amostras (metástase, tumor primário e sangue), quando disponível, foi obtido do mesmo paciente, sendo que essa casuística não possui informações sobre terapia oferecida aos pacientes. As amostras geradas pelo equipamento HiSeq X Ten (Illumina) foram obtidas no formato “.fastq” utilizando o SRA Run Selector, disponível no NCBI. Considerando esse e outros estudos relatados anteriormente, aproximadamente 35 *terabytes* de dados foram avaliados neste projeto.

3.2.6 Processamento e análise dos dados

Para a análise proteômica, os arquivos no formato “.mzML” foram processados e analisados em relação à quantificação relativa de proteínas utilizando o *software* MaxQuant (v.2.1.3.0) seguindo os parâmetros de processamento especificados em cada estudo utilizado. Todas as proteínas foram identificadas com um grau de confiança $\geq 95\%$ utilizando-se de busca na base de dados revisados Human UniProt (<https://www.uniprot.org/taxonomy/9606>) usando o mecanismo Andromeda (Cox et al., 2011). As proteínas identificadas como “reverse”, “only identified by site” e “potential contaminant” foram removidas dos arquivos finais “proteinGroups.txt” gerados após o processamento. Assim, considerando as matrizes com as abundâncias relativas obtidas tanto pela análise dos dois estudos públicos quanto pelo repositório CPTAC, as proteínas foram filtradas de modo a manter apenas aquelas com valor de abundância em, pelo menos 90% das amostras, considerando diferenciais aquelas que apresentaram $|\log_2 \text{fold-change}| \geq 1$ e $p < 0,05$ após correção de testes múltiplos pela taxa de falso-positivo (FDR, do inglês *False Discovery Rate*) pelo pacote limma (v.3.50.0, R/Bioconductor) (Ritchie

et al. 2015). Uma análise de componentes principais (PCA, do inglês *principal component analysis*) também foi realizada para auxiliar na escolha dos alvos com maior variabilidade entre as comparações, considerando aqueles que apresentaram o valor do componente principal 1 ($PC1 \geq 0,80$). Quando analisado o perfil global de expressão dos alvos entre suas respectivas amostras, a distribuição dos dados e a homogeneidade das variâncias foram verificadas a fim de atribuir o teste estatístico adequado para cada comparação, aplicando o teste de Mann-Whitney (não-paramétrico) ou teste-T não-pareado (paramétrico).

O processamento dos dados de RNA-seq foi feito utilizando um pipeline construído com o Nextflow (<https://www.nextflow.io>) e docker (<https://www.docker.com>). Os arquivos “.fastq” tiveram a qualidade avaliada com o programa FastQC e os adaptadores foram removidos com o programa Trim Galore. O alinhamento foi feito com o programa STAR e a quantificação pela ferramenta Salmon. Os dados foram então integrados em uma única matriz de contagem de *reads*, filtrados para manter valores de contagem ≥ 1 em, pelo menos, 90% das amostras, e corrigidos com a função ComBat do pacote sva (v.3.42.0, R/Bioconductor) considerando efeitos de lote conhecidos entre os conjuntos amostrais (Leek et al. 2012). Após essa etapa, os valores dispostos na matriz única foram normalizados e transformados pelo pacote DESeq2 (v.1.34.0, R/Bioconductor) (Love et al. 2014). Os dados de *microarray*, uma vez devidamente integrados, foram normalizados pelo método *quantile* e também corrigidos considerando os efeitos de lote conhecidos pelo pacote limma (Ritchie et al. 2015). A identificação de alvos diferencialmente expressos tanto para os dados de RNA-seq quanto *microarray* seguiu os mesmos critérios propostos anteriormente (Ritchie et al. 2015). Todas essas análises foram conduzidas no ambiente R utilizando pacotes base e outros que auxiliam na manipulação e visualização dos dados como tidyverse (v.1.3.1), plyr (v.1.8.6), dplyr (v.1.0.7), ggplot2 (v.3.3.5) e ComplexHeatmap (v.2.10.0) (Gu et al. 2016).

A identificação da atividade de vias de sinalização foi conduzida com o pacote PROGENy (do inglês *Pathway responsive genes for activity inference from gene expression*) (Schubert et al. 2018). O programa utiliza assinaturas para a inferência da atividade de vias de sinalização representadas em uma matriz de expressão por meio do *z-score* e *p*-valor ($<0,05$). Dessa maneira, vias identificadas com *z-score* ≤ -2 são consideradas inativadas e vias com *z-score* ≥ 2 são consideradas ativas. Também foi verificado o perfil imunológico em dados de *microarray*, RNA-seq e proteômica com o

pacote *immunedeconv* (Sturm et al. 2020). Esse método de deconvolução também investiga matrizes de expressão por meio de assinaturas de células B, macrófagos M1, macrófagos M2, monócitos, neutrófilos, células NK, células T CD4⁺ (não-regulatórias), células T CD8⁺, células T reguladoras (Tregs) e células dendríticas mielóides. Foram consideradas significativas as populações identificadas com $p < 0,05$.

Para a análise dos dados de *single-cell RNA-seq*, os *reads* foram alinhados contra o genoma de referência humano GRCh38 e as contagens do UMI (do inglês *Unique Molecular Identifier*) foram resumidas para cada célula. As matrizes geradas foram convertidas em um objeto Seurat pelo pacote Seurat (v.4.1.1), disponível para o R, e então filtradas para (1) remover células com baixo número de genes detectados (< 500), (2) remover células para as quais o número total de UMI (após transformação em $\log_{10}$) não estava dentro dos três desvios-padrão da média, (3) remover as células que apresentaram um número anormal de genes devido a um viés da técnica, (4) descartar células em que a proporção da contagem de UMI atribuída a genes mitocondriais foi superior a 15%. Após essas etapas, 128.616 células permaneceram no estudo, sendo 49.224 de tumores primários do cólon, 52.559 de metástases hepáticas e 26.833 de células mononucleares do sangue periférico.

Posteriormente, foi utilizada a função FindClusters em 30 componentes principais para realizar o agrupamento e a anotação das populações. A etapa de predição e classificação das células foi feita com o pacote scAnnotatR (Nguyen e Griss 2022) que investiga 34 populações por meio de marcadores incluindo CD3D, CD3G, TRAC (células T), CD19, CD79A e MS4A1 (células B), IGHG1, IGHA1, MZB1 e CD79A (plasmócitos), CD68, CD163, CD14 e LYZ (macrófagos), entre outras. Todas as etapas para visualização dos resultados e exploração dos dados foram conduzidas no ambiente R.

3.3 PROTEÔMICA DE METÁSTASES HEPÁTICAS INICIALMENTE IRRESSECÁVEIS DO CÂNCER DE CÓLON

Esta etapa do estudo de caráter retrospectiva foi avaliada e aprovada pelo Comitê de Ética em Pesquisa (CEP) do A.C.Camargo Cancer Center, sob o nº 3236/22 (Anexo 1). As amostras de metástase hepática foram obtidas a partir de fragmentos coletados por biópsia percutânea com agulha grossa (*core biopsy*) guiada por ultrassonografia ou

tomografia computadorizada e foram fixadas em formalina e emblocadas em parafina de acordo com os protocolos pré-estabelecidos pelo *Core Facility* de Histologia e Imuno-histoquímica do A.C.Camargo Cancer Center.

3.3.1 Seleção das amostras

A seleção de pacientes foi realizada entre março e julho de 2022 a partir do *software* institucional RECRUIT, uma base de dados que permite a rápida identificação de casos por critérios clínicos pré-estabelecidos. As buscas foram feitas utilizando os termos “biópsia de fígado”, “metástase”, “biópsia hepática”, “biópsia de lesão hepática” e “biópsia de nódulos hepáticos” e os filtros “[Diagnóstico CID-10] – Colon” e “[TNM 6a edição] - Cólon e Reto EC IV”, considerando suas respectivas combinações. Paralelamente, a enfermeira de pesquisa Bruna Elisa Catin de Sousa, que atua ativamente na área de tumores colorretais, forneceu uma lista de potenciais casos para serem integrados ao rastreamento inicial. Os 617 casos inicialmente levantados foram pré-filtrados com o auxílio do *software* de registros médicos H2TC, também institucional, com o objetivo de manter apenas aqueles pacientes que possuem biópsia hepática de metástases não-ressecáveis de câncer de cólon. Por fim, a seleção final dos pacientes foi feita com os *softwares* H2TC e Tasy considerando os casos que: a) realizaram o procedimento minimamente invasivo antes de receber qualquer tratamento sistêmico; b) estavam programados para receber quimioterapia para doença metastática; c) apresentavam dados clínicos e de evolução suficientes para permitir a avaliação da resposta ao tratamento (**Figura 3**).

As lâminas com cortes histológicos corados pela técnica de Hematoxilina-Eosina (H&E) e os blocos de parafina dos 46 pacientes incluídos no estudo foram solicitados ao *Core Facility* de Histologia e Imuno-histoquímica do A.C.Camargo Cancer Center. Nessa etapa, os pacientes que não apresentavam blocos de parafina disponíveis na instituição foram excluídos. Todas as amostras foram avaliadas pelo patologista com expertise na área, Dr. Gabriel Oliveira dos Santos, que é responsável pelo grupo da Patologia Gastrointestinal do A.C. Camargo Cancer Center. Após revisão das lâminas, foram descartados os casos que apresentaram amostra insatisfatória ou material insuficiente para as análises (**Figura 3**). Por fim os dados clínicos dos 15 casos selecionados foram coletados após análise dos prontuários disponíveis no H2TC e Tasy e inseridos na plataforma REDCap (do inglês *Research Eletronic Data Capture*) da instituição.

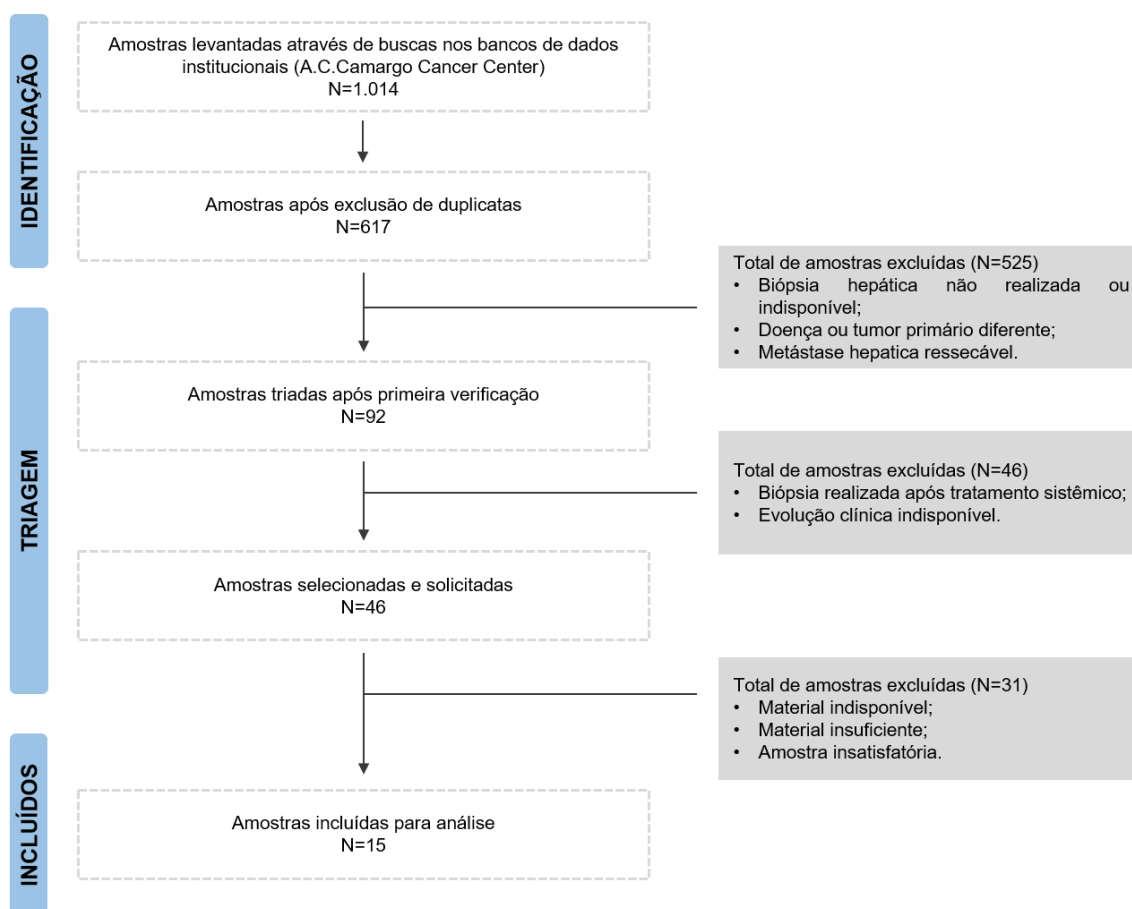


Figura 3 – Fluxograma representando as etapas de seleção das amostras FFPE rastreadas na instituição.

3.3.2 Casuística

Ao todo 30 amostras FFPE foram investigadas, sendo 15 de metástases hepáticas e 15 de tecidos adjacentes não-neoplásicos, coletadas de pacientes diagnosticados com câncer de cólon e metástases hepáticas irresssecáveis anteriormente ao tratamento sistêmico e que estavam programados para receber terapia padrão de primeira linha com quimioterapia e/ou mAb. A resposta à terapia foi avaliada com uso do RECIST 1.1 (do inglês *Response Evaluation Criteria in Solid Tumors*) utilizando métodos de imagem a cada 3 meses para verificar o comportamento da doença diante o tratamento sistêmico ofertado (Schwartz et al. 2016). Durante todo o período de avaliação, que foi terminado em óbito ou perda de seguimento, todos os pacientes não responderam completamente à terapia quando comparadas com as características anteriores ao início do tratamento, sendo: resposta parcial (PR) definida como uma diminuição de, pelo menos, 30% na soma do maior diâmetro de todas as lesões-alvo; doença progressiva (PD) definida como um

aumento de, pelo menos, 20% na soma do maior diâmetro de todas as lesões-alvo ou o aparecimento de uma ou mais lesões; ou doença estável (SD), que engloba avaliações que não apresentam nem um aumento suficiente para classificar como uma PD nem uma diminuição que qualifique uma PR (Schwartz et al. 2016).

3.3.3 Preparação das amostras para análise proteômica

Os blocos de parafina selecionados foram seccionados em micrótomo (RM2255, Leica) de modo a obter duas secções na espessura de 4-5 μm que foram preparadas em lâminas para microscopia (7109-WHITE, Precision) referentes a cada caso. Em seguida, o material foi sobreposto às lâminas coradas por H&E demarcadas pelo patologista colaborador para permitir a delimitação das áreas de interesse e macrodissecção com lâminas de bisturi estéreis. As áreas correspondentes ao tecido tumoral e não-tumoral adjacente de cada paciente foram coletadas separadamente em microtubos tratados com metanol e armazenadas a -20°C até o momento de desparafinização.

A retirada da parafina das amostras submetidas à macrodissecção foi feita com três adições de 500 μL xilol, cada uma envolvendo uma centrifugação a 15.000 rcf por 10 minutos e posterior descarte do sobrenadante. Posteriormente, as amostras foram hidratadas duas vezes com 500 μL de etanol absoluto e, em seguida, com soluções de etanol em concentrações decrescentes (90%, 70% e 50%), centrifugando a 15.000 rcf por 15 minutos e descartando o sobrenadante a cada etapa. Os solventes residuais foram removidos com um concentrador de amostras a vácuo (SpeedVac, Thermo Fisher Scientific) e as amostras foram ressuspensas em 30 μL de bicarbonato de amônio 50 mM e armazenadas a 4°C *overnight*. Após esse período, o sobrenadante foi recuperado em novos microtubos tratados com metanol e as amostras foram novamente secas a vácuo e armazenadas a -80°C para futura extração de proteínas.

3.3.4 Extração e digestão de proteínas

Os experimentos foram conduzidos no Laboratório Multiusuários de Proteômica, Metabolômica e Lipidômica Max Feffer, localizado no Departamento de Genética na Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo (ESALQ-USP, Piracicaba/SP). Todas as amostras foram ressuspensas vigorosamente com 400 μL de tampão para extração e solubilização de proteínas contendo ureia 7M, tiourea 2M, DTT 10 mM e Triton X-100 a 0,01%. Após centrifugação a 16.000 rcf por

15 minutos a 4°C, o sobrenadante contendo as proteínas extraídas foi recuperado e as amostras foram dessalinizadas com o auxílio de filtros Amicon® Ultra 3KDa (Merck Millipore), substituindo o tampão utilizado por bicarbonato de amônio 50 mM. Nessa etapa, as colunas foram previamente hidratadas com água ultrapura e centrifugadas a 14.000 rcf por 10 minutos. Posteriormente, o sobrenadante de cada amostra foi adicionado às colunas, centrifugando a 14.000 rcf por 15 minutos ou até atingir volume de 100 µL. O líquido percolado foi descartado e 300 µL de bicarbonato de amônio 50 mM foram adicionados, seguindo novo ciclo de centrifugação. Essa etapa foi repetida 5 vezes. Na última etapa, o volume final próximo de 100 µL restante na coluna de dessalinização foi vertido em novo microtubo tratado com metanol e as amostras foram secas a vácuo.

A concentração das amostras complexas de proteínas foi determinada pelo método de Bradford (Bradford, 1976), verificando a absorbância das amostras em um comprimento de onda de 595 nm. Dois microgramas (2 µg) de cada amostra em um volume final de 10 µL completado com bicarbonato de amônio 50 mM foram desnaturados separadamente com 2,5 µL de solução de RapiGest™ SF a 0,2% (Waters) a 80°C, por 15 minutos; reduzidos com 0,625 µL de solução-estoque de ditiotreitol 100 mM (GE Healthcare) à fim de chegar a molaridade final de 5 mM a 60°C, por 30 minutos; e alquilados com 0,66 µL de solução-estoque de iodoacetamida 300 mM (GE Healthcare) à fim de chegar a molaridade final de 15 mM, por 30 minutos no escuro à temperatura ambiente. As amostras foram então digeridas com 4 µL de tripsina (*Sequencing Grade Modified Trypsin*, Promega) a uma proporção de 1:100 (w/w) enzima:proteína, por 16 horas, a 37°C. Após digestão triptica, 1 µL de solução de ácido trifluoroacético (TFA) 5% foi adicionado às amostras para hidrolisar a ação do RapiGest™ SF utilizando a proporção de 0,4/1 (TFA 5%/RapiGest™ SF), incubando a 37°C por 90 minutos. Em seguida, a mistura de peptídeos foi centrifugada a 16.000 rcf por 30 minutos a 4°C e o sobrenadante foi transferido para novo microtubo e seco a vácuo. Os peptídeos resultantes da digestão foram ressuspensos e dessalinizados com auxílio de ponteiros ZipTip com resina C₁₈ (octadesilsilano) (Merck Millipore) de acordo com as instruções do fabricante. Finalmente, as amostras foram novamente secas a vácuo e mantidas a -80°C até análise por LC/MS. Para injeção direta das amostras no equipamento, o volume final de 10 µL foi obtido pela adição de solução aquosa contendo 0,1% de ácido fórmico (AF) de modo a obter uma concentração final de 200 ng.µL⁻¹.

3.3.5 Análise do perfil proteômico por LC/MS

As análises de LC/MS foram realizadas com o auxílio do sistema de nanocromatografia líquida de alta eficiência nanoElute® (Bruker Daltonics, Germany) acoplado on-line a um espectrômetro de massas do tipo timsTOF Pro (do inglês *hybrid trapped ion mobility spectrometry-quadrupole time-of-flight mass spectrometer*) (Bruker Daltonics, Germany) com fonte de ionização por ESI (CaptiveSpray, Bruker Daltonics, Germany). Uma alíquota (1 µL) de amostra equivalente a 200 ng de peptídeos digeridos foi injetada em uma coluna analítica *Aurora 2* (IonOpticks, Austrália) com fase estacionária à base de C₁₈ (partículas de 1,6 µm, 250 mm de comprimento por 75 µm de diâmetro) e estes foram separados a uma taxa de fluxo de 250 nL.min⁻¹ com um gradiente típico de fase reversa (Solvente A: 0,1% de AF, 99,9% de água Milli-Q; Solvente B: 0,1% de AF, 99,9% de acetonitrila). A temperatura da coluna foi mantida a 50°C e a corrida cromatográfica foi de 120 minutos (2% a 15% do Solvente B durante 60 minutos, elevada a 25% aos 90 minutos, elevada a 37% aos 100 minutos, elevada a 95% aos 110 minutos e finalmente 95% por 10 minutos para lavagem da coluna). A temperatura da linha de transferência capilar de íons foi ajustada para 180°C e o acúmulo de íons por 123 milissegundos, e então a separação por mobilidade foi obtida com uma rampa de potencial de entrada de -160V a -20V dentro do período estabelecido.

Durante a aquisição de dados, para habilitar o método de acúmulo paralelo à fragmentação dos íons (PASEF, do inglês *parallel accumulation serial fragmentation*), as informações de precursores de *m/z* e mobilidade foram primeiro derivadas de um experimento de *full scan* considerando um intervalo de *m/z* entre 100 e 1700. Precursores monocarregados foram excluídos por sua posição no plano *m/z*-íon de mobilidade, e os precursores que atingiram o valor alvo de 20.000 unidades arbitrárias foram dinamicamente excluídos por 0,4 minuto. O modo operacional do timsTOF Pro e do método PASEF foram controlados e sincronizados com o auxílio do *software* OtofControl (v.5.1) de controle instrumental (Bruker Daltonics, Germany).

3.3.6 Parâmetros de processamento e busca em banco de dados públicos

O processamento dos dados, a identificação de proteínas e as análises de quantificação relativa foram realizadas utilizando-se do *software* MaxQuant (v.2.1.3.0). Os parâmetros de processamento incluíram: carbamidometilação da cisteína como modificação fixa; oxidação da metionina e acetilação da região N-terminal como

modificações variáveis; dois possíveis erros de clivagem pela enzima proteolítica tripsina; mínimo de um peptídeo único para identificação de proteínas; e uma taxa máxima de falso-positivo de 1%. Os espectros MS/MS foram pesquisados no banco de dados revisado Human UniProt (<https://www.uniprot.org/taxonomy/9606>) usando o mecanismo de busca Andromeda (Cox et al., 2011) e todas as proteínas foram identificadas com um grau de confiança $\geq 95\%$. Os arquivos denominados “proteinGroups.txt”, gerados pelo processamento do MaxQuant, foram utilizados para as análises e todas as proteínas listadas como “reverse”, “only identified by site” e “potential contaminant” foram removidas. Somente as proteínas identificadas em $>70\%$ das amostras foram consideradas para a análise estatística, que foi realizada pelo pacote limma (Ritchie et al., 2015). Para as análises de inferência foram imputados valores médios para os dados ausentes utilizando o pacote impute (v.1.70.0), disponível no Bioconductor. Foram consideradas proteínas diferencialmente abundantes aquelas com $FC \geq 2$ e ≤ -2 e $p < 0,05$ após comparações entre os grupos amostrais. Além disso, as proteínas diferenciais também foram analisadas em relação à variabilidade entre a comparação de metástases e tecido adjacente não-neoplásico, selecionando os alvos com $PC1 \geq 0,80$.

4 RESULTADOS

4.1 CARACTERÍSTICAS GERAIS DOS DADOS PÚBLICOS DE EXPRESSÃO GÊNICA

As buscas nos principais repositórios públicos visaram a identificação e seleção do maior número possível de estudos contendo dados de expressão gênica de amostras de câncer de cólon e metástases hepáticas. Todos os estudos foram analisados manualmente para identificar se esses atendiam aos critérios de seleção propostos, separando-os de acordo com a plataforma utilizada e informações clínicas disponíveis. Uma lista final de 58 estudos foi obtida. Estudos com um tamanho amostral relativamente baixo ($N < 10$ para tumores primários) não foram excluídos, mas foram destacados na planilha final, com exceção do GSE40367 que possui amostras de tumores primários de 7 pacientes pareados com metástases hepáticas. Em suma, dentre os estudos selecionados ($N=58$), 10 deles foram realizados por RNA-seq e 48 por *microarray*.

Os estudos alocados em diferentes plataformas (customizadas, Affymetrix, Agilent, Illumina ou RNA-seq) foram subdivididos em grupos de acordo com os dados clínicos disponíveis. Aqueles com avaliação da resposta à terapia foram classificados como “Tratamento” e os que apresentaram dados de sobrevida foram classificados como “Sobrevida” – os conjuntos de dados contendo ambas as informações foram denominados “Tratamento+Sobrevida”. O restante dos estudos que continham apenas amostras de tumores primários foram classificados como “TNM”. Os estudos com metástases hepáticas foram classificados como “Metástases” ou “Metástases+[...]” dependendo do grupo em que as amostras de tumores primários se encaixavam. A **Tabela 1** resume essas informações para 20 estudos e as planilhas que contém todas as informações sobre o processo de rastreamento e seleção dos estudos foram disponibilizadas no Google Drive (<<https://drive.google.com>>).

Tabela 1 – Dados de expressão gênica de 20 estudos classificados de acordo com a plataforma utilizada e dados clínicos disponíveis.

ID	Data	Grupo	Amostra	NT	Tumor	Tipo	Localização	Estágio	Tratamento	Sobrevida	Meta	Plataforma
GSE101481 ^a	2019	TNM	Fresca	NA	58	Cólon	Sim	Sim	Não	Não	1	GPL23782
GSE25071 ^a	2011	TNM	Fresca	4	46	CCR	Sim	Sim	Não	Não	NA	GPL2986
E-TIGR-16 ^a	2005	TNM	Fresca	NA	36	Cólon	Não	Sim	Não	Não	NA	TIGR
E-MTAB-333 ^a	2011	Tratamento	Fresca	NA	27	Cólon	Sim	Sim	Sim	Não	NA	A-MEXP-503
GSE41258 ^b	2012	Metástases+Sobrevida	Fresca	45	185	CCR	Sim	Sim	Não	Sim	47	GPL96
GSE131418 ^b	2019	Metástases+Sobrevida	Fresca	NA	142	CCR	Sim	Sim	Sim	Sim	146	GPL15048
GSE40967 ^b	2013	Sobrevida	Fresca	19	566	CCR	Sim	Sim	Não	Sim	NA	GPL570
GSE62080 ^b	2014	Tratamento	Fresca	NA	21	Cólon	Não	Sim	Sim	Não	NA	GPL570
GSE72970 ^b	2017	Tratamento+sobrevida	Fresca	NA	86	Cólon	Sim	Sim	Sim	Sim	NA	GPL570
GSE81423 ^c	2017	Metástases	FFPE	NA	NA	NA	NA	NA	NA	NA	63	GPL14951
GSE14297 ^c	2010	Metástases+TNM	Fresca	7	18	CCR	Sim	Sim	Não	Não	18	GPL6370
GSE53127 ^c	2013	Tratamento+sobrevida	FFPE	NA	18	Cólon	Não	Sim	Sim	Sim	NA	GPL8432
GSE63596 ^d	2016	TNM	Fresca	NA	15	CCR	Sim	Sim	Não	Não	NA	GPL17077
GSE146480 ^d	2020	TNM	Fresca	NA	10	Cólon	Sim	Sim	Não	Não	NA	GPL16699
GSE104645 ^d	2018	Tratamento+sobrevida	FFPE	NA	193	Cólon	Sim	Sim	Sim	Não	NA	GPL6480
GSE144259 ^e	2020	Metástases+TNM	Fresca	3	3	Cólon	Não	Sim	Não	Não	3	GPL11154
GSE147749 ^e	2020	Metástases+TNM	Fresca	NA	2	Cólon	Não	Sim	Não	Não	2	GPL20301
GSE92914 ^e	2018	Metástases+TNM	Fresca	1	8	Cólon	Sim	Sim	Não	Não	3	GPL18573
GSE145432 ^e	2020	Metástases	Fresca	NA	12	CCR	Não	Sim	Não	Não	10	GPL18573
E-MTAB-8412 ^e	2020	TNM	Fresca	NA	8	Cólon	Não	Não	Não	Não	NA	GPL20301

^a Estudos realizados com plataformas de *microarray* customizadas. ^b Estudos realizados com plataformas de *microarray* da Affymetrix. ^c Estudos realizados com plataformas de *microarray* da Illumina. ^d Estudos realizados com plataformas de *microarray* da Agilent. ^e Estudos realizados com a técnica de RNA-seq com plataformas da Illumina. TNM: sistema baseado na avaliação do tamanho e/ou extensão do tumor primário, acometimento de linfonodos regionais e metástases à distância; FFPE: amostras fixadas em formalina e embebidas em parafina; NT: tecido não-neoplásico; CCR: câncer colorretal; Meta: metástase hepática.

4.2 DADOS CLÍNICO-PATOLÓGICOS DOS CASOS INCLUÍDOS NA ANÁLISE PROTEÔMICA

O acesso aos prontuários eletrônicos permitiu a obtenção dos dados clínico-patológicos e de tratamento dos 15 casos selecionados, incluindo informações acerca da idade ao diagnóstico, localização da doença primária, status mutacional dos principais genes preditores de eficácia terapêutica, quimioterapia ofertada como primeira linha de tratamento e o tempo de seguimento clínico na instituição. Esses dados, assim como outras informações, estão descritos na **Tabela 2**.

Tabela 2 – Resumo das características clínicas e histopatológicas dos 15 pacientes com metástases hepáticas irresssecáveis do câncer de cólon incluídos nesse trabalho.

Variáveis	Número de pacientes N=15
Sexo (%)	
Feminino	9 (60%)
Masculino	6 (40%)
Idade ao diagnóstico do tumor primário (anos)	
Mediana	63
Variação	50-84
Idade ao diagnóstico da metástase (anos)	
Mediana	63
Variação	50-88
Localização (%)	
Ceco (C18.0)	1 (6,7%)
Ascendente (C18.2)	3 (20%)
Transverso (C18.4)	2 (13,3%)
Sigmoide (C18.7)	6 (40%)
Não especificado (C18.9)	3 (20%)
Metástase (%)	
Sincrônica	14 (93,3%)
Metacrônica	1 (6,7%)
Status mutacional do KRAS no tumor primário (%)	
Selvagem	6 (40%)
Mutado	2 (13,3%)
Não avaliado	7 (46,7%)
Status mutacional do KRAS na metástase (%)	
Selvagem	3 (20%)
Mutado	4 (26,7%)
Inconclusivo	1 (6,6%)
Não avaliado	7 (46,7%)
Status mutacional do NRAS no tumor primário (%)	
Selvagem	6 (40%)
Mutado	1 (6,7%)
Não avaliado	8 (53,3%)
Status mutacional do NRAS na metástase (%)	
Selvagem	3 (20%)
Mutado	0 (0%)
Inconclusivo	1 (6,7%)
Não avaliado	11 (73,3%)

Con/Tabela 2 – Resumo das características clínicas e histopatológicas dos 15 pacientes com metástases hepáticas irressecáveis do câncer de cólon incluídos nesse trabalho.

Variáveis	Número de pacientes N=15
Status mutacional do <i>BRAF</i> no tumor primário (%)	
Selvagem	6 (40%)
Mutado	0 (0%)
Não avaliado	9 (60%)
Status mutacional do <i>BRAF</i> na metástase (%)	
Selvagem	2 (13,3%)
Mutado	0 (0%)
Inconclusivo	1 (6,7%)
Não avaliado	12 (80%)
Tratamento de primeira linha (%)	
FOLFOX	4 (26,7%)
FOLFOX+mAb	3 (20%)
FOLFIRI	4 (26,7%)
FOLFIRI+mAb	3 (20%)
FOLFIRINOX+mAb	1 (6,6%)
Seguimento (meses)	
Mediana	20
Variação	3-39
Tempo de progressão (%)	
≤6 meses	7 (46,7%)
>6 meses	7 (46,7%)
Inconclusivo (perda de seguimento)	1 (6,6%)
Status vital (%)	
Vivo com a doença	2 (13,3%)
Óbito	10 (66,7%)
Inconclusivo (perda de seguimento)	3 (20%)

mAb: anticorpo monoclonal; FOLFOX: 5-fluorouracil, leucovorina e oxaliplatina; FOLFIRI: 5-fluorouracil, leucovorina e irinotecano; FOLFIRINOX: 5-fluorouracil, leucovorina, oxaliplatina e irinotecano.

A maioria dos pacientes incluídos na casuística eram do sexo feminino e todas apresentaram metástase hepática sincrônica, com uma idade média ao diagnóstico de 65 anos. Apenas um caso do sexo masculino desenvolveu a doença secundária após quatro anos do diagnóstico e tratamento cirúrgico do tumor primário (metástase metacrônica) – excluindo esse caso, esses indivíduos apresentavam uma média de 59 anos durante o manejo inicial. Em relação à localização do tumor primário, quase metade dos pacientes desenvolveram tumores no cólon sigmoide (40%), seguido do cólon ascendente correspondendo a 20% dos casos avaliados. Além disso, em 3 casos (20%) não foi possível especificar a localização exata do tumor primário, assim classificando-os como C18.9 (neoplasia maligna do cólon, sem outras especificações).

Considerando o regime quimioterápico ofertado como primeira linha de tratamento, 4 pacientes receberam FOLFOX e outros 4 receberam FOLFIRI (26,7% dos casos, respectivamente). Seis pacientes receberam combinações associadas ao bevacizumabe, sendo 2 casos administrados com FOLFOX (33,3%), 3 com FOLFIRI

(50%) e 1 em conjunto com FOLFIRINOX (16,7%). O mAb cetuximabe foi apenas administrado em um paciente da casuística em combinação com FOLFOX. Apesar de todos os pacientes terem sido classificados como não-respondedores à quimioterapia de acordo com a evolução clínica disponibilizada nos prontuários eletrônicos, 6 pacientes apresentaram PR à primeira linha de tratamento, 3 casos não apresentaram alterações significativas nos exames de imagem, se enquadrando em SD, 5 manifestaram PD, e o paciente que recebeu FOLFIRINOX demonstrou redução significativa de todas as lesões, desenvolvendo PD hepática após 7 meses dessa constatação.

Os pacientes também foram avaliados quanto ao status mutacional dos genes *KRAS*, *NRAS* e *BRAF*. Para avaliação do *KRAS*, o material proveniente do tumor primário foi usado em 8 casos, sendo que um paciente apresentou mutação no códon 12 (G12D) e outro exibia mutação no códon 61 (Q61L) do gene. Outros 7 pacientes foram avaliados a partir da amostra de biópsia hepática e 43% desses (3 casos) possuíam mutação G12D e 14% (1 caso) mutação no códon 13 (G13D). O tumor primário foi utilizado para verificar mutações em *NRAS* em 6 casos, e apenas um apresentou mutação G12D. Por sua vez, o status selvagem foi identificado no tecido metastático de 3 pacientes – essa avaliação não foi realizada em 5 pacientes. Finalmente, 8 pacientes exibiram *BRAF* no status selvagem, em que 5 casos foram avaliados a partir do material do cólon e 2 da metástase hepática – não houve avaliação em 7 casos.

Quanto ao tempo compreendido entre o primeiro ciclo de tratamento e a primeira constatação de PD, 7 pacientes (46,7%) apresentaram progressão nos primeiros 6 meses após o início da terapia de primeira linha. Desses, cerca de 57% (4 casos) mostraram resultados desfavoráveis no quinto mês. Em contrapartida, outros 7 pacientes (46,7%) desenvolveram PD após esse período estabelecido, mostrando uma variação entre 8 a 30 meses (mediana de 12 meses). Para um caso não foi possível obter esse dado por perda de seguimento, porém o paciente foi encaminhado aos cuidados paliativos após 7 meses do início do tratamento. Por fim, em relação ao tempo de seguimento e status vital, 66,7% dos pacientes (10 casos) vieram a óbito pela doença, contando com um tempo médio de *follow-up* de 18 meses. Dos casos remanescentes, 2 (13,3%) permanecem vivos com a doença e 3 (20%) perderam seguimento após um período de aproximadamente 16 meses a partir da entrada no A.C.Camargo Cancer Center.

4.3 PERFIL TRANSCRIPTÔMICO E PROTEÔMICO GLOBAL DAS METÁSTASES

Foram integradas 1.478 amostras de câncer de cólon primário, tanto M0 quanto M1, tecido adjacente não-neoplásico do cólon, tecido saudável do fígado e cólon (provindas do GTEx) e metástases inicialmente ressecáveis do fígado. O PCA da **Figura 4A** mostra a projeção inicial das amostras antes da normalização e correção dos efeitos de lote. É possível observar que, mesmo considerando um plano bidimensional e um PC1 de aproximadamente 35%, não é possível distinguir claramente cada conjunto de dados. A **Figura 4B** apresenta o PCA após normalização e correção dos efeitos de lote utilizando a função `Combat_seq` do pacote `sva`. Esse resultado representa com maior clareza a disposição esperada dos grupos avaliados, principalmente devido a heterogeneidade dos dados. Após o tratamento dos dados foi feito um *heatmap* para observarmos o agrupamento das amostras de acordo com o perfil de expressão. Na **Figura 4C** observamos uma tendência de agrupamento dos conjuntos de acordo com o perfil molecular e em concordância com a projeção do PCA.

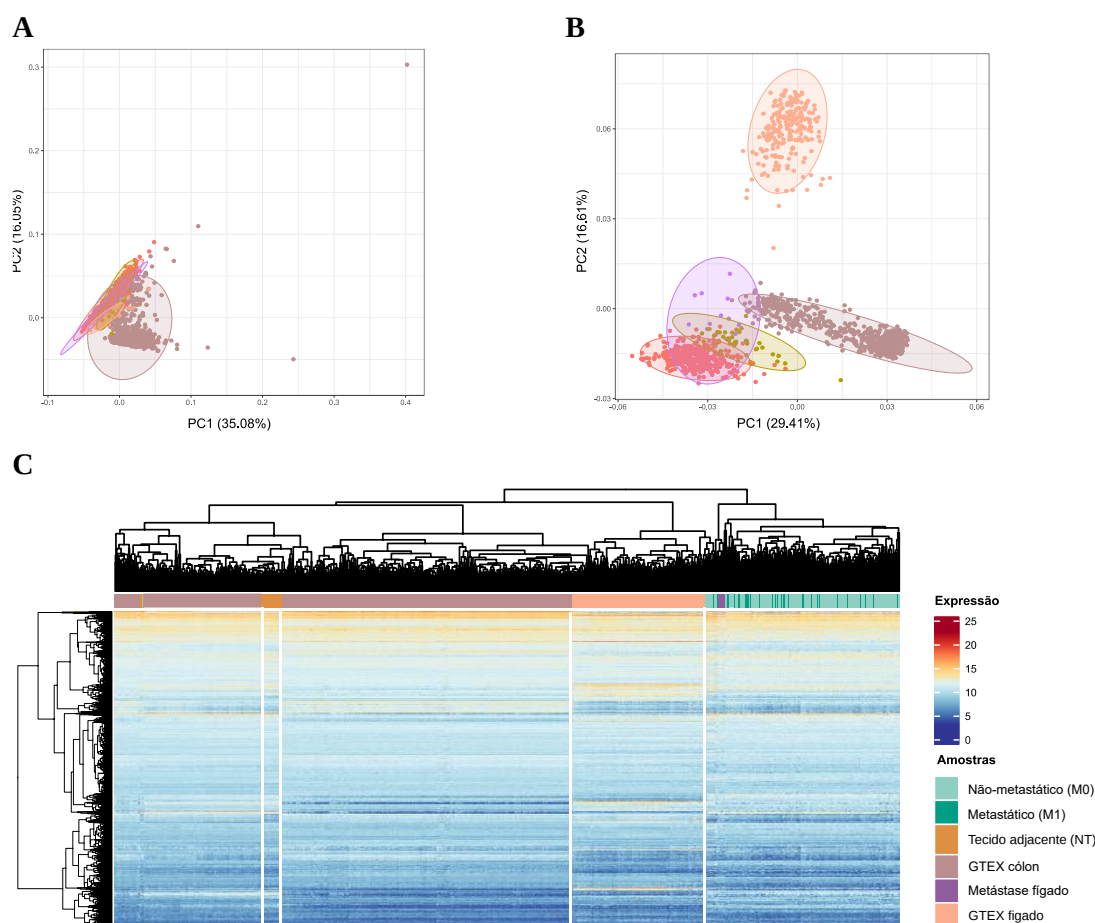


Figura 4 - PCA e *heatmap* das amostras de RNA-seq investigadas na meta-análise. **(A)** PCA dos grupos de amostras incluídos antes da normalização e correção do efeito de lote. **(B)** PCA dos grupos de amostras após a normalização e correção do efeito de lote. **(C)** Perfil de expressão do conjunto de 1.478 amostras avaliadas na meta-análise. O *heatmap* mostra o agrupamento dos 6 grupos de amostras incluídos e avaliados na meta-análise: tumores primários de cólon M0 e M1, tecidos adjacentes não-neoplásicos (NT), tecidos saudáveis do cólon e fígado obtidos do banco de dados GTEx e metástases hepáticas ressecáveis. Vermelho: aumento de expressão; Azul: diminuição de expressão.

Após o processamento e normalização dos dados foram feitas comparações para identificação de genes diferencialmente expressos entre os conjuntos investigados. Foram realizadas todas as comparações de acordo com as combinações possíveis, mas optamos por apresentar apenas assinaturas relacionadas com a comparação entre metástases hepáticas e tumores M1 devido a enorme quantidade de dados gerados. Foram obtidos 193 genes diferenciais, dos quais a **Tabela 3** apresenta os 40 principais, sendo 20 com aumento de expressão e 20 com diminuição de expressão. Para cada alvo selecionado foi apresentado também o *fold-change* (FC) e p-valor obtido para as outras comparações. Devido a grande quantidade de genes selecionados, também apresentamos o valor do PC1 obtido na PCA considerando cada comparação. Essa análise tem como objetivo auxiliar a escolha dos melhores genes para as próximas etapas.

Tabela 3 - Representação dos 40 principais genes diferencialmente expressos entre metástases hepáticas e CCR_{M1}.

	LM vs M1			LM vs M0		LM vs GTEx _{cólon}		LM vs GTEx _{fígado}		LM vs N _{positivo}		LM vs N _{negativo}	
Expressão aumentada													
Genes	FC	adj.Pval	PCA _{PC1}	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval
<i>APOB</i>	4,83	1,21E-03	0,31	4,46	6,04E-04	25,8	8,34E-15	-3,10	1,30E-02	4,54	1,21E-04	4,66	9,46E-04
<i>APOA1#</i>	4,75	1,01E-04	0,80	2,10	3,72E-03	24,15	8,66E-20	3,12	2,75E-03	3,73	6,61E-04	3,97	6,65E-04
<i>ITIH4</i>	4,57	1,47E-03	0,32	3,92	8,86E-04	---	---	---	---	3,85	5,75E-04	3,95	1,07E-04
<i>ADH1B</i>	4,37	5,51E-04	0,66	4,57	5,82E-05	---	---	---	---	4,05	7,54E-04	4,74	3,14E-05
<i>HP</i>	4,24	4,10E-04	0,02	4,42	2,53E-05	11,66	4,58E-07	---	---	4,10	1,80E-03	4,51	1,26E-04
<i>VTN</i>	3,99	1,82E-05	0,46	3,92	9,98E-04	6,24	4,70E-05	---	---	3,74	1,89E-03	4,02	4,00E-02
<i>TNFR2#</i>	3,89	4,63E-05	0,88	-2,66	2,35E-05	31,43	3,11E-13	12,18	1,22E-19	3,17	5,39E-04	3,86	8,45E-06
<i>PRELP</i>	3,84	5,47E-04	0,77	-4,16	1,66E-04	---	---	8,46	1,02E-05	3,32	4,08E-06	4,42	1,43E-03
<i>MASP1</i>	3,66	5,96E-05	0,66	3,65	8,93E-06	---	---	---	---	3,24	1,62E-04	3,80	3,35E-06
<i>VCL</i>	3,60	8,38E-05	0,81	3,76	5,00E-06	3,13	7,33E-04	13,07	6,38E-06	3,27	1,39E-04	4,04	7,34E-03
<i>CUX2</i>	3,27	4,76E-05	0,15	3,30	6,53E-03	3,64	3,03E-04	---	---	3,29	1,91E-05	3,29	6,23E-03
<i>ITIH2#</i>	3,24	1,95E-05	0,81	3,40	6,28E-03	7,65	1,78E-05	2,25	5,41E-08	3,19	7,01E-06	3,54	2,18E-04
<i>MDH2#</i>	3,23	1,83E-04	0,87	-2,34	1,76E-05	3,33	3,32E-02	5,55	2,99E-04	3,23	7,30E-05	3,35	1,78E-05
<i>CDH2#</i>	3,21	1,96E-04	0,89	3,37	1,39E-05	2,02	2,12E-02	7,93	1,28E-04	2,83	5,18E-04	3,64	2,30E-04
<i>TGFB1#</i>	3,20	3,02E-05	0,83	-2,14	6,22E-06	2,83	1,33E-02	8,22	2,08E-06	2,9	6,73E-05	3,21	3,80E-03
<i>A1BG</i>	3,13	1,53E-03	0,52	2,91	6,86E-03	---	---	---	---	2,87	3,39E-03	2,94	6,38E-02
<i>NFKB1#</i>	3,12	5,94E-05	0,91	-2,81	6,37E-05	2,84	4,10E-02	2,88	3,88E-05	2,68	2,79E-04	2,87	4,61E-05
<i>FTCD</i>	3,08	8,45E-04	0,42	3,12	5,68E-05	3,02	1,39E-02	5,01	1,67E-04	2,83	1,49E-05	3,22	2,61E-03
<i>C3#</i>	3,04	2,43E-05	0,84	2,83	1,47E-05	2,91	1,44E-02	12,73	2,19E-09	2,65	1,10E-04	2,99	4,51E-02
<i>EGFR#</i>	3,03	5,58E-04	0,91	-2,10	8,90E-05	2,24	3,19E-04	2,20	4,42E-02	2,67	1,34E-03	3,33	2,20E-05
Expressão diminuída													
Genes	FC	adj.Pval	PCA _{PC1}	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval
<i>SMIM3</i>	-5,80	9,02E-05	0,39	-6,28	1,15E-05	-103,83	1,45E-126	-25,51	1,77E-65	-6,17	3,99E-05	-6,18	1,42E-06
<i>JCHAIN</i>	-3,52	2,64E-03	0,66	-3,98	1,35E-04	-6,24	1,36E-02	---	---	-3,63	1,04E-03	-3,92	1,91E-04
<i>TRIL</i>	-3,01	2,31E-04	0,67	-3,53	8,57E-04	-12,76	2,24E-24	-3,39	3,93E-06	-3,35	1,55E-03	-3,47	2,29E-02
<i>MMP3</i>	-3,15	5,20E-05	0,86	-3,43	9,25E-05	-3,07	2,75E-02	---	---	-2,95	7,28E-05	-3,50	7,21E-03
<i>ADAMDEC1</i>	-3,03	1,76E-03	0,44	-3,33	1,22E-04	-6,62	1,81E-03	---	---	-3,31	2,52E-04	-3,18	2,47E-04
<i>GOLGA6L9</i>	-3,12	1,63E-04	0,34	-3,20	2,77E-05	-13,48	6,56E-82	-7,82	1,54E-51	-3,19	6,46E-03	-3,22	2,98E-03
<i>IGHV3-23</i>	-2,73	4,43E-03	0,47	3,17	1,98E-04	-3,56	3,89E-02	---	---	-2,80	2,03E-03	-3,06	3,79E-04
<i>CLCA4</i>	-2,85	2,56E-02	0,30	-2,99	8,27E-03	-9,77	3,33E-03	---	---	-2,67	2,83E-02	-2,96	9,04E-03
<i>MTX1P1#</i>	-2,51	3,62E-05	0,87	2,72	2,45E-03	-2,71	5,47E-22	-2,27	1,68E-14	-2,71	1,04E-04	-2,71	2,96E-03

Con/Tabela 3 - Representação dos 40 principais genes diferencialmente expressos entre metástases hepáticas e CCR_{M1}.

Genes	LM vs M1			LM vs M0		LM vs GTEx _{cólon}		LM vs GTEx _{fígado}		LM vs N _{positivo}		LM vs N _{negativo}	
	FC	adj.Pval	PCA _{PC1}	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval	FC	adj.Pval
<i>ZDHHC20</i>	-2,92	5,49E-03	0,45	-2,63	1,36E-03	-3,41	9,05E-07	-2,70	8,46E-05	-2,72	7,08E-04	-2,58	5,53E-03
<i>LRRC15</i>	-2,62	3,95E-05	0,28	-2,60	6,44E-06	-3,60	1,60E-03	---	---	-2,77	3,58E-06	-2,39	3,92E-05
<i>DIO2</i>	-2,72	4,86E-03	0,29	2,60	8,97E-03	-2,75	1,68E-03	---	---	-2,66	1,87E-03	-2,44	1,11E-03
<i>OLFM4</i>	-2,47	4,20E-02	0,51	-2,59	1,60E-02	---	---	---	---	---	---	-2,69	1,35E-02
<i>PPA1</i>	-2,65	8,24E-02	0,65	-2,58	2,53E-02	-2,30	1,69E-03	-3,37	2,69E-06	-2,35	1,39E-03	-2,61	2,95E-03
<i>TXNDC9</i>	-2,86	2,64E-02	0,10	2,51	1,36E-04	-3,28	1,42E-05	-3,94	4,80E-07	-2,66	4,56E-02	-2,47	3,96E-02
<i>PDIA6</i>	-2,55	1,29E-02	0,66	-2,45	7,29E-04	-2,28	7,56E-04	-4,06	3,62E-09	-2,35	3,81E-02	-2,48	5,99E-03
<i>UMAD1</i>	-2,57	1,90E-03	0,61	2,31	2,48E-02	-3,22	7,01E-04	-3,06	1,36E-03	-2,48	9,15E-04	-2,21	6,93E-02
<i>VBP1</i>	-2,65	1,06E-03	0,47	-2,31	3,90E-03	-3,50	1,99E-05	-2,08	1,74E-02	-2,33	9,93E-04	-2,32	4,28E-02
<i>GJB2</i>	-2,60	7,85E-03	0,14	-2,14	5,36E-05	---	---	-7,07	5,79E-08	-2,13	1,37E-04	-2,10	8,70E-05
<i>CCZ1P1</i>	-2,52	4,42E-03	0,41	2,08	3,08E-05	---	---	-5,28	1,21E-06	-2,20	2,63E-05	-2,04	5,31E-05

Considerando a assinatura de 193 genes diferencialmente expressos encontrada entre metástases hepáticas (LM) e tumores metastáticos (M1), responsável pelo agrupamento de ambos os conjuntos investigados, esses resultados também foram comparados quando observado amostras não-metastáticas (M0), linfonodo positivo e negativo, e de tecido normal do cólon e fígado. FC: *fold-change*; adj.Pval: p-valor ajustado pela taxa de falso-positivo (FDR, do inglês *False Discovery Rate*); N: linfonodo; # representa PC1 $\geq$ 0,80 para todas as comparações; “---” representa não-identificação de genes de acordo com os critérios estatísticos estabelecidos.

Foi possível observar resultados interessantes, como por exemplo a identificação dos genes *TNFR2* (FC=3,89; $p=4,63E-05$), *MDH2* (FC=3,23; $p=1,83E-04$), *TGFB1* (FC=3,20; $p=3,02E-05$), *EGFR* (FC=3,03; $p=5,58E-04$), *NFKB1* (FC=3,12; $p=5,94E-05$), *APOA1* (FC=4,75; $p=1,01E-04$), *ITIH2* (FC=3,24; $p=1,95E-05$), *CDH2* (FC=3,21; $p=1,96E-04$), *C3* (FC=3,04; $p=2,43E-05$) e *MTX1P1* (FC=-2,51; $p=3,62E-05$), que obtiveram valor de $PC1 \geq 0,80$ na projeção realizada com o PCA, justificando um possível papel de destaque para esses candidatos. Entretanto, não foram observadas diferenças significativas entre os resultados obtidos dos apresentados quando considerado amostras com ou sem disseminação regional das células tumorais (linfonodo positivo e negativo, respectivamente), além disso, os genes destacados apresentaram a mesma direção do sinal do FC entre as comparações.

Partindo do conjunto dos 20 principais genes diferencialmente expressos com expressão aumentada, uma análise de enriquecimento para vias de sinalização do KEGG utilizando o programa EnrichR foi realizada (Kuleshov et al. 2016). Considerando p-valor ajustado para $FDR < 0,05$, 17 vias significativas foram identificadas e representadas na **Tabela 4**.

Tabela 4 - Enriquecimento para vias de sinalização com os 20 principais genes diferencialmente expressos com expressão aumentada identificados na comparação entre metástases hepáticas e CCR_{M1} utilizando a base de dados do KEGG.

Vias de sinalização	N.genes	adj.Pval	Genes
Adesão focal	3 201	1,00E-03	<i>VTN, EGFR, VCL</i>
Proteoglicanos no câncer	3 205	1,06E-03	<i>VTN, TGFB1, EGFR</i>
Doença inflamatória intestinal	2 65	1,90E-03	<i>TGFB1, NFKB1</i>
Sinalização de células epiteliais na infecção por <i>Helicobacter pylori</i>	2 70	2,20E-03	<i>NFKB1, EGFR</i>
Junção aderente	2 71	2,27E-03	<i>EGFR, VCL</i>
Via de sinalização MAPK	3 294	2,98E-03	<i>TGFB1, NFKB1, EGFR</i>
Câncer colorretal	2 86	3,30E-03	<i>TGFB1, EGFR</i>
Expressão de PD-L1 e via de checkpoint PD-1 no câncer	2 89	3,53E-03	<i>NFKB1, EGFR</i>
Infecção por <i>Staphylococcus aureus</i>	2 95	4,01E-03	<i>C3, MASP1</i>
Via de sinalização AGE-RAGE em complicações diabéticas	2 100	4,43E-03	<i>TGFB1, NFKB1</i>
Via de sinalização PI3K/AKT	3 354	5,01E-03	<i>VTN, NFKB1, EGFR</i>
Diferenciação de células Th17	2 107	5,06E-03	<i>TGFB1, NFKB1</i>
Via de sinalização HIF-1	2 109	5,24E-03	<i>NFKB1, EGFR</i>
Via de sinalização FoxO	2 131	7,49E-03	<i>TGFB1, EGFR</i>
Formação de armadilhas extracelulares de neutrófilos	2 189	1,51E-02	<i>C3, NFKB1</i>
Via de sinalização Ras	2 232	2,22E-02	<i>NFKB1, EGFR</i>
Rede imune intestinal para produção de IgA	1 48	4,69E-02	<i>TGFB1</i>

Na coluna “N.genes” a barra separa o número de genes alvos identificados em cada módulo e a quantidade de genes que representam toda a via. adj.Pval: p-valor ajustado pela taxa de falso-positivo (FDR, do inglês *False Discovery Rate*).

É possível observar importantes vias de sinalização que foram enriquecidas com os alvos selecionados, incluindo CCR, via da MAPK, PD-1, PI3K/AKT, entre outras. Também se destacam alguns genes que foram mais representativos nessa análise, como por exemplo *TGFB1* (N=8 vias), *NFKB1* (N=9 vias) e *EGFR* (N=11 vias). A estratégia de realizar a análise de enriquecimento com maior enfoque nos principais alvos selecionados levou a diminuição na quantidade de vias identificadas e no número de genes presentes em cada um dos módulos enriquecidos. Quando essa mesma análise foi realizada com os 20 principais genes com expressão diminuída não foram identificadas vias de sinalização significativas.

4.4 PERFIL GLOBAL E PROTEÍNAS DIFERENCIALMENTE ABUNDANTES ENTRE METÁSTASES IRRESSECÁVEIS E TECIDOS ADJACENTES

Com relação às análises de LC-MS realizadas com as amostras de metástases hepáticas irressecáveis e tecidos pareados não-neoplásicos, 1.788 proteínas foram detectadas e 526 se destacaram com $FC \geq 2$ e ≤ -2 e $p < 0,05$. Entretanto, optamos por considerar o p-valor ajustado por $FDR < 0,05$ na tentativa de remover possíveis *outliers* que poderiam influenciar na escolha dos biomarcadores. Dessa maneira, apenas 46 proteínas se mantiveram na análise. De modo geral, uma das possíveis justificativas para o baixo número global de alvos selecionados foi a menor quantidade de proteínas que foram utilizadas no experimento para algumas amostras, uma vez que a escassez de material biológico disponível nas biópsias impossibilitou a obtenção de um número maior de proteínas. Entretanto, esse fator não influenciou na representatividade das proteínas em cada amostra, uma vez que foram mantidas apenas aquelas presentes em toda a casuística, além do processo de normalização dos dados. A **Figura 5** apresenta um *heatmap* com o perfil das 526 proteínas detectadas, onde é possível observar o agrupamento esperado entre metástases hepáticas e entre amostras adjacentes não-neoplásicas.

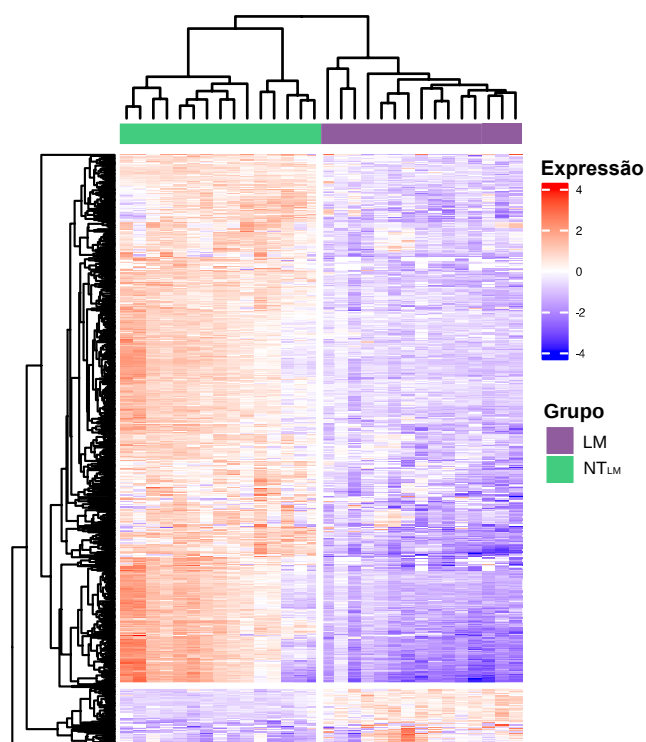


Figura 5 - Perfil global das 526 proteínas selecionadas após comparação entre metástases hepáticas inicialmente irresssecáveis e tecidos adjacentes não-neoplásicos. Observa-se o agrupamento esperado entre amostras metastáticas e entre não-neoplásicas. LM: metástase hepática; NT_{LM}: tecido pareado não-neoplásico adjacente ao tumor. Vermelho: aumento de abundância; Azul: diminuição de abundância.

A **Tabela 5** apresenta as principais proteínas identificadas com p-valor ajustado por FDR, sendo 26 com aumento da abundância e 20 com redução. Entre aquelas com diminuição de abundância nas metástases, o menor FC foi observado para a proteína CXCL14 (FC=-4,52; p=4,58E-03), acompanhada por CLCA1 (FC=-4,42; p=1,07E-03), TFF1 (FC=-4,40; p=3,22E-02) e PLAC8 (FC=-4,40; p=1,87E-02). Entre as proteínas com aumento de abundância, YWHAZ (FC=22,55; p=1,56E-02), AHNAK (FC=21,17; p=2,70E-04), HNRNPH1 (FC=20,31; p=3,79E-03), PCBP1 (FC=16,47; p=3,25E-02) e TGFB1 (FC=14,08; p=3,48E-02) apresentaram os maiores valores de FC. Quando os resultados foram projetados com a PCA, os principais candidatos que se mantiveram com PC1 $\geq$ 0,80 foram YWHAZ, PCBP1, TGFB1, TOP2A, TKT, CDH2, ENO1, APOA1, TNFR2 e EGFR. Outra característica observada durante a seleção dos principais candidatos foi a maior concentração de proteínas com abundância diminuída (aproximadamente 80% do total) entre as 526 selecionadas previamente após comparação entre metástases e tecidos adjacentes. Apesar disso, a grande maioria delas não foi detectada após ajuste do p-valor por FDR.

Tabela 5 - Proteínas diferencialmente abundantes selecionadas após a comparação entre metástases inicialmente irresssecáveis do fígado e tecido adjacente.

Proteínas	ILM vs NT			RLM vs NT			ILM vs RLM			ILM vs M1			ILM vs M0		
	FC	adj.Pval	PCA	FC	adj.Pval	PCA	FC	adj.Pval	PCA	FC	adj.Pval	PCA	FC	adj.Pval	PCA
YWHAZ#	22,55	1,56E-02	0,85	-2,08	5,03E-01	0,88	24,63	2,59E-03	0,82	3,53	1,20E-01	0,82	3,43	4,97E-03	0,89
AHNAK	21,17	2,70E-04	0,67	-2,02	7,46E-01	0,35	23,19	4,86E-03	0,65	3,49	9,62E-02	0,61	2,85	6,64E-02	0,60
HNRNPH1	20,31	3,79E-03	0,06	2,29	1,63E-02	0,55	18,02	8,25E-04	0,14	3,81	6,75E-03	0,05	3,36	5,33E-04	0,10
PCBP1#	16,47	3,25E-02	0,85	2,26	1,77E-02	0,85	14,21	3,70E-05	0,86	4,25	2,42E-04	0,88	3,90	8,49E-07	0,85
TGFB1#	14,08	3,48E-02	0,88	2,01	9,06E-01	0,80	12,07	4,25E-03	0,85	3,43	1,02E-01	0,81	3,29	4,80E-03	0,84
DDX39A	4,84	2,93E-04	0,67	---	---	---	---	---	---	---	---	---	---	---	---
ITIH2#	4,69	3,55E-03	0,85	-2,00	8,85E-01	0,86	6,69	6,64E-01	0,86	2,58	8,49E-01	0,90	2,28	7,33E-01	0,80
TOP2A#	4,48	3,36E-02	0,82	2,10	5,37E-01	0,85	2,38	4,42E-01	0,87	-2,41	8,83E-01	0,81	2,01	9,75E-01	0,88
TKT	4,33	1,60E-02	0,70	-2,08	2,56E-01	0,87	6,41	2,46E-04	0,88	3,77	1,40E-03	0,80	3,11	5,35E-04	0,82
CDH2#	4,28	3,72E-02	0,82	2,29	4,64E-01	0,82	1,99	9,70E-01	0,85	2,35	9,64E-01	0,91	2,31	9,16E-01	0,83
ENO1#	4,16	2,96E-02	0,84	2,51	6,74E-03	0,81	1,65	3,17E-05	0,80	5,77	2,67E-04	0,86	4,71	5,87E-05	0,85
HIST1H2BL	4,05	6,45E-03	0,70	---	---	---	---	---	---	---	---	---	---	---	---
HNRNPA1	3,99	3,95E-02	0,33	2,00	9,32E-01	0,55	1,99	2,50E-01	0,59	3,00	2,92E-01	0,65	2,79	5,65E-02	0,37
APOA1#	3,83	2,99E-02	0,87	2,14	2,80E-01	0,82	1,69	2,99E-01	0,89	2,69	6,34E-01	0,87	2,67	8,51E-02	0,84
TNFR2#	3,62	3,60E-02	0,81	-2,04	6,50E-01	0,88	5,66	1,39E-01	0,84	3,25	2,44E-01	0,81	3,06	3,80E-02	0,80
EGFR#	3,60	3,26E-02	0,82	-1,99	9,06E-01	0,80	5,59	1,72E-01	0,84	3,36	1,67E-01	0,88	3,02	3,88E-02	0,82
HIST2H3PS2	3,50	3,73E-02	0,64	2,35	2,70E-02	0,69	1,15	8,63E-01	0,26	-2,54	8,49E-01	0,39	2,18	8,63E-01	0,09
EEF2	3,34	1,72E-02	0,41	-2,06	4,77E-01	0,70	5,40	8,14E-01	0,11	2,51	8,27E-01	0,22	2,08	9,35E-01	0,55
SERPINA1	3,31	3,71E-02	0,30	-2,21	1,92E-01	0,73	5,52	3,92E-01	0,71	-2,30	9,70E-01	0,47	-2,18	8,77E-01	0,07
HIST1H2AJ	3,25	3,41E-02	0,76	2,18	1,47E-01	0,76	1,07	4,25E-03	0,58	2,99	2,62E-01	0,08	3,02	8,12E-03	0,07
TPM4	3,11	1,81E-02	0,46	-2,08	8,95E-01	0,56	5,19	2,13E-01	0,56	3,00	8,52E-01	0,78	3,42	4,69E-01	0,54
HSP90AA1	2,72	3,05E-02	0,77	2,14	1,73E-01	0,51	0,58	1,07E-01	0,72	3,25	5,23E-02	0,80	2,78	1,47E-02	0,64
RPS3	2,65	4,86E-02	0,16	-2,01	6,50E-01	0,51	4,66	5,18E-01	0,46	2,91	2,44E-01	0,76	2,48	1,55E-01	0,78
MDH2#	2,64	4,83E-02	0,80	-2,19	1,15E-01	0,84	4,83	6,51E-01	0,84	3,00	3,36E-01	0,82	2,22	7,33E-01	0,85
HIST1H4A	2,53	4,00E-02	0,75	---	---	0,41	---	---	0,42	---	---	0,49	---	---	0,05
EEF1A1	2,43	2,04E-02	0,66	1,97	9,89E-01	0,82	0,46	3,00E-01	0,89	2,58	6,79E-01	0,82	2,32	3,59E-01	0,84
CXCL14#	-4,52	4,58E-03	0,81	-1,96	9,32E-01	0,84	-2,56	4,55E-01	0,91	2,62	6,75E-01	0,85	2,68	5,58E-02	0,82
CLCA1	-4,42	1,07E-03	0,21	2,26	2,00E-02	0,14	-6,68	4,43E-02	0,06	2,75	4,61E-01	0,54	2,43	2,59E-01	0,49
TFF1	-4,40	3,22E-02	0,39	-2,19	3,82E-01	0,60	-2,21	9,51E-01	0,82	2,50	8,80E-01	0,83	2,13	9,41E-01	0,55
PLAC8	-4,40	1,87E-02	0,44	-1,95	9,56E-01	0,04	-2,45	2,30E-02	0,76	2,91	2,87E-01	0,52	2,79	2,57E-02	0,65

Con/Tabela 5 - Proteínas diferencialmente abundantes selecionadas após a comparação entre metástases inicialmente irressecáveis do fígado e tecido adjacente.

Proteínas	ILM vs NT			RLM vs NT			ILM vs RLM			ILM vs M1			ILM vs M0		
	FC	adj.Pval	PCA	FC	adj.Pval	PCA	FC	adj.Pval	PCA	FC	adj.Pval	PCA	FC	adj.Pval	PCA
PLEC	-3,97	4,77E-02	0,49	---	---	0,21	---	---	0,78	---	---	0,17	---	---	0,67
SPARCL1	-3,93	1,26E-02	0,30	-2,43	6,13E-05	0,25	-1,50	8,55E-01	0,39	2,31	9,44E-01	0,54	-1,98	9,81E-01	0,14
P2RY14	-3,65	3,61E-02	0,73	-2,09	6,32E-01	0,01	-1,56	4,71E-01	0,52	3,28	4,61E-01	0,38	2,72	3,99E-01	0,54
MAT2A	-3,50	9,61E-03	0,68	-2,15	3,15E-01	0,73	-1,35	5,67E-01	0,41	-2,33	9,48E-01	0,40	1,97	9,89E-01	0,62
PLA2G2A#	-3,49	3,62E-02	0,88	-2,12	9,03E-02	0,80	-1,37	4,64E-01	0,84	2,55	6,97E-01	0,82	2,13	7,50E-01	0,82
FCGBP	-3,40	1,54E-02	0,56	1,99	9,80E-01	0,33	-5,39	8,60E-01	0,42	2,38	9,37E-01	0,83	2,01	9,81E-01	0,09
TMEM158	-3,24	4,53E-02	0,57	-1,97	9,06E-01	0,20	-1,27	7,51E-01	0,35	-2,60	7,16E-01	0,29	-1,98	9,84E-01	0,14
FBLN1	-3,22	2,20E-02	0,77	2,07	4,53E-01	0,02	-5,29	5,47E-01	0,78	2,26	9,91E-01	0,73	-1,98	9,81E-01	0,69
ACTG2	-3,21	2,34E-02	0,31	-2,30	2,55E-02	0,11	-0,91	7,02E-01	0,72	2,74	6,73E-01	0,63	2,33	5,91E-01	0,12
CEACAM5	-3,17	1,05E-02	0,71	-2,08	5,50E-01	0,48	-1,09	8,63E-02	0,87	3,46	1,88E-01	0,45	2,76	2,24E-01	0,42
SPINK4	-3,07	2,52E-02	0,72	-2,00	8,21E-01	0,84	-1,07	5,20E-01	0,74	2,61	7,16E-01	0,76	2,29	5,69E-01	0,82
PGM5	-3,05	4,24E-03	0,52	-2,61	2,56E-07	0,35	-0,44	6,08E-01	0,25	2,99	2,87E-01	0,43	2,12	8,77E-01	0,01
MYL9	-3,03	4,28E-05	0,58	---	---	0,65	---	---	0,09	---	---	0,05	---	---	0,75
CAV1#	-2,71	2,37E-02	0,84	3,86	3,78E-01	0,84	-6,57	8,14E-01	0,83	-3,30	8,80E-01	0,88	1,97	9,97E-01	0,81
LPAR1	-2,37	3,00E-02	0,71	-2,28	5,03E-04	0,67	-0,09	7,60E-01	0,46	-2,38	8,53E-01	0,29	-1,99	9,75E-01	0,64
SMTN	-2,03	2,29E-03	0,37	2,56	8,76E-01	0,44	-4,59	9,23E-01	0,43	3,01	9,48E-01	0,37	-2,95	9,69E-01	0,66

ILM: metástase hepática irressecável; RLM: metástase hepática ressecável; M0: tumores primários não-metastáticos; M1: tumores primários metastáticos; NT: tecido adjacente não-neoplásico; FC: *fold-change*; adj.Pval: p-valor ajustado pela taxa de falso-positivo (FDR, do inglês *False Discovery Rate*); # representa $PC1 \geq 0,80$ para todas as comparações; “---” representa não-identificação de genes de acordo com os critérios estatísticos estabelecidos.

Ainda com relação à **Tabela 5**, é interessante observar entre as proteínas com aumento de expressão a presença de componentes de histonas como HIST1H2BL, HIST2H3PS2, HIST1H2AJ e HIST1H4A, com atividades relacionadas a estrutura de nucleossomos e fibra cromossômica. Apesar de não terem sido selecionadas de acordo com a projeção do PCA, todas apresentaram variabilidade considerável com valores médios da PC1 próximos a 0,7.

Partindo do conjunto de proteínas com níveis elevados de abundância nas metástases hepáticas, o programa EnrichR foi utilizado para análise do enriquecimento global (seleção por FDR<0,05) e 21 vias de sinalização associadas foram identificadas. Na **Tabela 6** é possível observar a identificação de vias relacionadas ao CCR ($p=5,55E-03$), HIF-1 ($p=1,98E-02$), PI3K/AKT ($p=4,77E-02$) e Ras ($p=1,89E-02$). Quando a mesma análise foi realizada com as proteínas com diminuição da abundância, 8 vias foram significativas, incluindo adesão focal ($p=1,70E-02$) e regulação do citoesqueleto de actina ($p=1,97E-02$).

Tabela 6 - Enriquecimento para vias de sinalização a partir das proteínas identificadas como diferencialmente abundantes na comparação entre metástases hepáticas e tecido adjacente não-neoplásico utilizando a base de dados do KEGG.

Vias de sinalização	N.proteínas	adj.Pval	Proteínas
Proteínas com abundância aumentada			
Câncer de pâncreas	2 76	4,37E-03	TGFB1, EGFR
Câncer colorretal	2 86	5,55E-03	TGFB1, EGFR
Câncer de próstata	2 97	7,01E-03	HSP90AA1, EGFR
Diferenciação de células Th17	2 107	8,48E-03	HSP90AA1, TGFB1
Via de sinalização HIF-1	2 109	8,78E-03	ENO1, EGFR
Via de sinalização PI3K/AKT	3 354	1,06E-02	HSP90AA1, YWHAZ, EGFR
Ciclo de célula	2 124	1,12E-02	TGFB1, YWHAZ
Via de sinalização da relaxina	2 129	1,21E-02	TGFB1, EGFR
Via de sinalização FoxO	2 131	1,25E-02	TGFB1, EGFR
Via de sinalização de estrogênio	2 137	1,36E-02	HSP90AA1, EGFR
Câncer de intestino	2 149	1,59E-02	TGFB1, EGFR
Spliceossoma	2 150	1,61E-02	PCBP1, HNRNPA1
Via de sinalização da ocitocina	2 154	1,70E-02	EEF2, EGFR
Hepatite C	2 157	1,76E-02	YWHAZ, EGFR
Hepatite B	2 162	1,87E-02	TGFB1, YWHAZ
Via de sinalização do Ras	2 163	1,89E-02	TGFB1, YWHAZ
Carcinoma hepatocelular	2 168	2,00E-02	TGFB1, EGFR
Proteoglicanos no câncer	2 205	2,89E-02	TGFB1, EGFR
Doença do coronavírus	2 232	3,63E-02	RPS3, EGFR
Carcinogênese química	2 239	3,83E-02	HSP90AA1, EGFR
Ciclo citrato (ciclo TCA)	1 30	3,83E-02	MDH2
Proteínas com abundância diminuída			
Vias de sinalização	N.proteínas	adj.Pval	Proteínas
Contração do músculo liso vascular	3 133	3,02E-04	PLA2G2A, MYL9, ACTG2
Secreção pancreática	2 102	4,61E-03	PLA2G2A, CLCA1
Adesão focal	2 201	1,70E-02	CAV1, MYL9

Con/Tabela 6 - Enriquecimento para vias de sinalização a partir das proteínas identificadas como diferencialmente abundantes na comparação entre metástases hepáticas e tecido adjacente não-neoplásico utilizando a base de dados do KEGG.

Vias de sinalização	N.proteínas	adj.Pval	Proteínas
Regulação do citoesqueleto de actina	2 218	1,97E-02	LPAR1, MYL9
Digestão e absorção de gorduras	1 43	4,22E-02	PLA2G2A
Interação neuroativa ligante-receptor	2 341	4,50E-02	P2RY14, LPAR1
Metabolismo de éter lipídico	1 49	4,79E-02	PLA2G2A
Metabolismo da cisteína e da metionina	1 50	4,89E-02	MAT2A

Foram investigadas 46 proteínas diferencialmente abundantes selecionadas na análise anterior quando considerado $FC \geq 2$ e ≤ -2 e $FDR < 0,05$. Na coluna “N.proteínas” a barra separa o número de proteínas identificadas em cada módulo e a quantidade de proteínas que representam toda a via. adj.Pval: p-valor ajustado pela taxa de falso-positivo (FDR, do inglês *False Discovery Rate*).

4.5 ATIVIDADE DA SINALIZAÇÃO CELULAR

Considerando que as análises de enriquecimento apenas identificam alvos com grande representação quando comparados a processos biológicos/vias de sinalização, utilizamos o programa PROGENy para analisar possíveis alterações na atividade de vias bem associadas à progressão tumoral e mecanismos de resistência. A **Figura 6** apresenta o perfil de ativação/inibição entre metástases hepáticas inicialmente irressecáveis e tecidos não-neoplásicos pareados. Foi observado o agrupamento entre as amostras de cada conjunto e formação de blocos de vias diferencialmente alteradas.

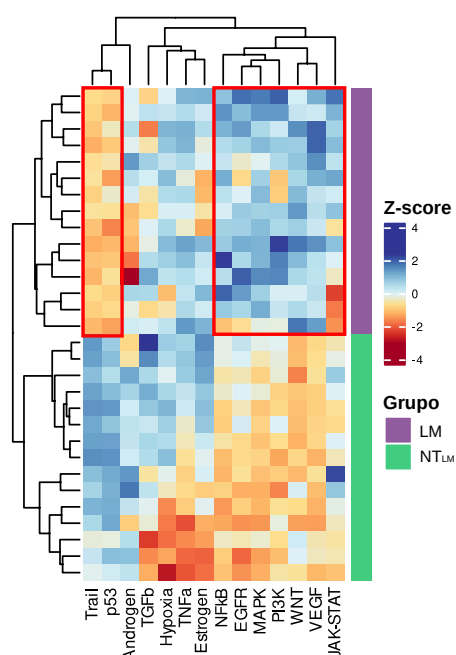


Figura 6 - *Heatmap* representando a atividade de sinalização entre metástases hepáticas irressecáveis e tecido não-neoplásico adjacente. O $z\text{-score} \geq 2$ define ativação da via (azul) e $z\text{-score} \leq -2$ define inibição da via (vermelho).

$score \leq -2$ define inibição (vermelho). LM: metástases hepáticas irressecáveis; NT_{LM}: tecido não-neoplásico adjacente.

Entre os resultados significativos destacam-se a ativação de JAK-STAT ($z-score=3,18$; $p=3,16E-03$), VEGF ($z-score=2,67$; $p=8,83E-03$), PI3K ($z-score=3,99$; $p=4,90E-03$), WNT ($z-score=3,22$; $p=2,24E-04$), MAPK ($z-score=2,85$; $p=5,27E-03$), EGFR ($z-score=2,49$; $p=1,43E-03$) e estrógeno ($z-score=3,43$; $p=6,66E-03$) em amostras metastáticas, enquanto p53 ($z-score=-2,29$; $p=1,99E-03$) e *trail* ($z-score=-3,29$; $p=7,70E-03$) apresentaram inativação nesse mesmo grupo (**Tabela 7**).

Tabela 7 - Perfil de ativação e inibição de vias de sinalização considerando os dados de proteômica obtidos de metástases hepáticas inicialmente irressecáveis e tecidos não-neoplásicos pareados.

Vias	LM vs. NT _{LM}	
	Score	P-valor
Andrógeno	0,87	3,85E-01
EGFR	2,49	1,43E-03
Estrógeno	3,43	6,66E-03
Hipóxia	0,28	2,50E-01
JAK-STAT	3,18	3,16E-03
MAPK	2,85	5,27E-03
NFkB	4,11	7,61E-03
p53	-2,29	1,99E-03
PI3K	3,99	4,90E-03
TGFb	0,32	1,90E-01
TNF- α	0,65	9,47E-01
<i>Trail</i>	-3,29	7,70E-03
VEGF	2,67	8,83E-03
WNT	3,22	2,24E-04

Todos os valores de ativação/inibição de vias foram representados em $z-score$. Dessa maneira, $z-score \geq 2$ indica ativação e $z-score \leq -2$ indica inativação. LM: metástase hepática; NT_{LM}: tecido adjacente não-neoplásico.

4.6 PERFIL IMUNOLÓGICO DAS AMOSTRAS INVESTIGADAS

A inferência do perfil imunológico das amostras de metástase hepática foi feita a partir do pacote *immunedeconv* (Sturm et al. 2019), que investiga a presença de células B, macrófagos M1, macrófagos M2, monócitos, neutrófilos, células NK, células T CD4+ (não-reguladoras), células T CD8+, células Tregs e células dendríticas mieloides. Para isso, uma lista de marcadores específicos é utilizada para cada população de células e a presença das mesmas é inferida nas amostras do estudo por meio de uma análise de deconvolução. Na **Figura 7** é possível observar uma diferença em relação a presença de neutrófilos ($p=2,52E-03$), células Tregs ($p=3,21E-06$) e macrófagos M1 ($p=1,33E-02$) em amostras de metástase hepática. Quando observadas as amostras não-neoplásicas

adjacentes, populações de células B foram significativamente enriquecidas ($p=2,87E-06$), além de maior presença de macrófagos M2 ($p=4,77E-02$).

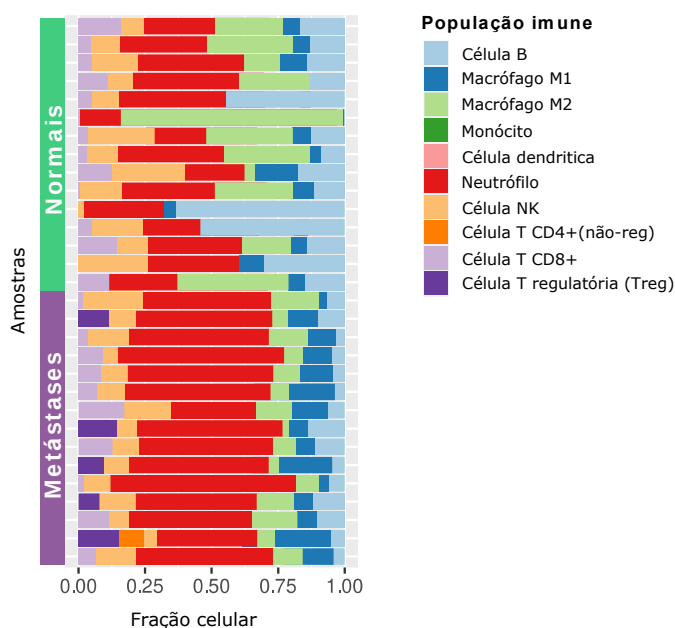


Figura 7 - Identificação de populações de células imunes nas amostras de metástase hepática e tecido não-neoplásico adjacente. A fração celular representa a % de células nas respectivas amostras.

4.7 COMPARAÇÃO ENTRE TRANSCRITOS E PROTEÍNAS E CONFIRMAÇÃO COM ESTUDOS INDEPENDENTES DA LITERATURA

Considerando os transcritos e proteínas significativos que foram selecionados após as comparações estatísticas destacam-se TGFB1, CDH2, APOA1, TNFR2, EGFR, MDH2 e ITIH2, que estavam presentes em ambas as análises e com $PC1 \geq 0,80$. Dessa maneira, esses resultados prosseguiram para análises posteriores com o objetivo de confirmar o papel de destaque desses alvos nas metástases. Além disso, também foram selecionadas as proteínas YWHAZ, PCBP1, TOP2A e ENO1 (abundância aumentada) e CXCL14, PLA2G2A e CAV1 (abundância diminuída) devido a grande variação ($PC1 \geq 0,80$) identificada no PCA. Uma das análises realizadas para validar esses achados foi a comparação com 364 amostras de metástases hepáticas ressecáveis e tumores primários de cólon obtidas a partir de cinco estudos independentes: GSE14297 (18 metástases e 18 tumores primários), GSE22834 (21 metástases e 31 tumores primários), GSE41568 (80 metástases e 39 tumores primários), GSE183984 (21 metástases e 84

tumores primários) e GSE190609 (6 metástases e 35 tumores primários). Os dados que foram utilizados nesta comparação são provenientes do perfil de expressão de transcritos obtidos tanto pela técnica de *microarray* quanto por *RNA-seq*, sugerindo uma interpretação paralela entre aumento de expressão de genes codificadores e maior regulação das proteínas e vice-versa. Na **Figura 8** é possível observar os resultados em forma de *boxplot* para os cinco estudos mostrando a confirmação da expressão de 12 dos 14 alvos investigados ($p < 0,05$). Para as proteínas PLA2G2A e CAV1 os resultados obtidos não foram significativos.

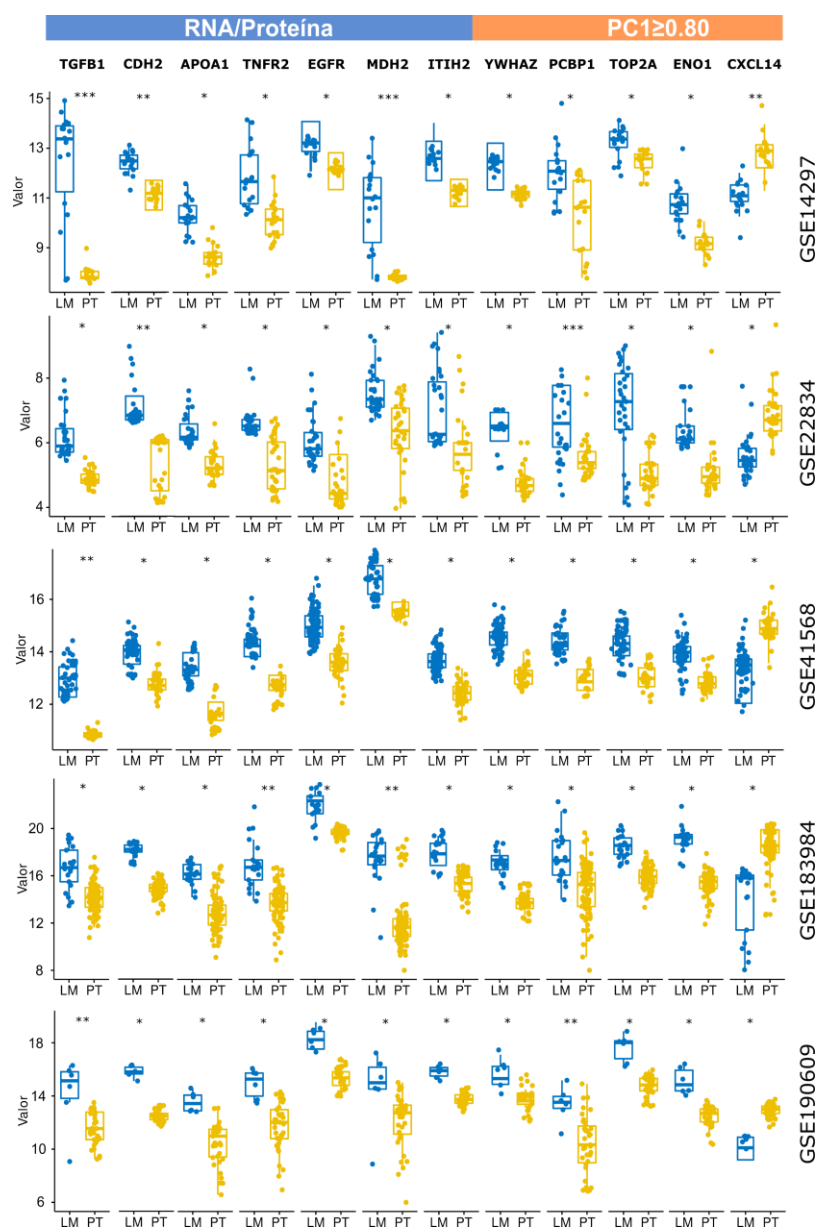


Figura 8 - Representação em *boxplot* da análise para confirmação do perfil de expressão das 14 proteínas selecionadas em cinco estudos independentes. Inicialmente foram investigados os perfis de expressão de 14 alvos, porém PLA2G2A e CAV1 não obtiveram resultados significativos. LM: metástases hepáticas ressecáveis; PT: tumores primários do cólon. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Buscando uma possível associação entre os 12 alvos que tiveram confirmação da expressão em metástases hepáticas e populações de células imunes, uma análise utilizando dados de *single-cell RNA-seq* foi feita a partir de dados do estudo GSE178318. Como descrito anteriormente, o estudo é composto por 15 amostras de seis pacientes, sendo 6 de metástases ressecáveis do fígado, 6 de câncer primário no cólon e 3 de sangue periférico coletado antes da cirurgia. Assim, os dados foram processados e analisados para a obtenção de agrupamentos (*clusters*) que são formados de acordo com a proximidade de células que compartilham características similares. A **Figura 9A** apresenta nas cores vermelha (câncer de cólon primário), verde (metástases hepáticas) e azul (células mononucleares do sangue periférico) todas as populações de células projetadas em duas dimensões pelo algoritmo UMAP. Após a identificação dos agrupamentos, as mesmas células foram anotadas para populações imunes de acordo com a presença de marcadores específicos. A **Figura 9B** apresenta todas as populações que foram identificadas com destaque para células T CD4, NK, plasmócitos e neutrófilos.

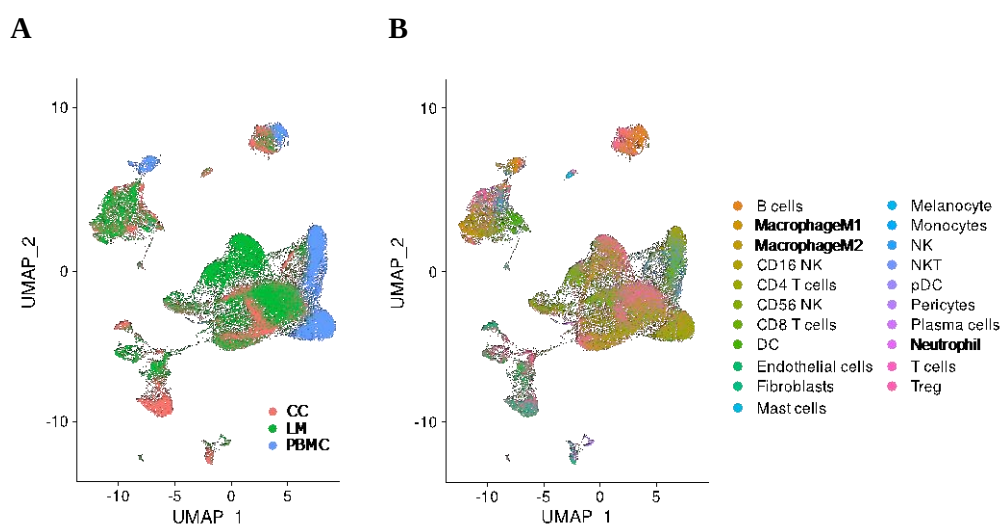


Figura 9 - Representação de amostras de metástases hepáticas, tumores primários do cólon e células mononucleares do sangue periférico obtidas por *single-cell RNA-seq*. **(A)** Projeção bidimensional utilizando UMAP de 128.616 células, sendo 49.224 CCs, 52.559 LMs e 26.833 PBMCs. Os grupos foram representados pelas cores vermelha, verde e azul, respectivamente. **(B)** Anotação dos agrupamentos obtidos para populações de células imunes através de marcadores específicos. CC: câncer primário do cólon; LM: metástases hepáticas ressecáveis; PBMC: células mononucleares do sangue periférico.

Posteriormente, todas as células de metástases hepáticas foram isoladas para a investigação da expressão dos 12 alvos selecionados anteriormente. Também foram marcadas isoladamente as populações imunes para observarmos possíveis sobreposições com a expressão das proteínas. Esses resultados foram apresentados na **Figura 10**.

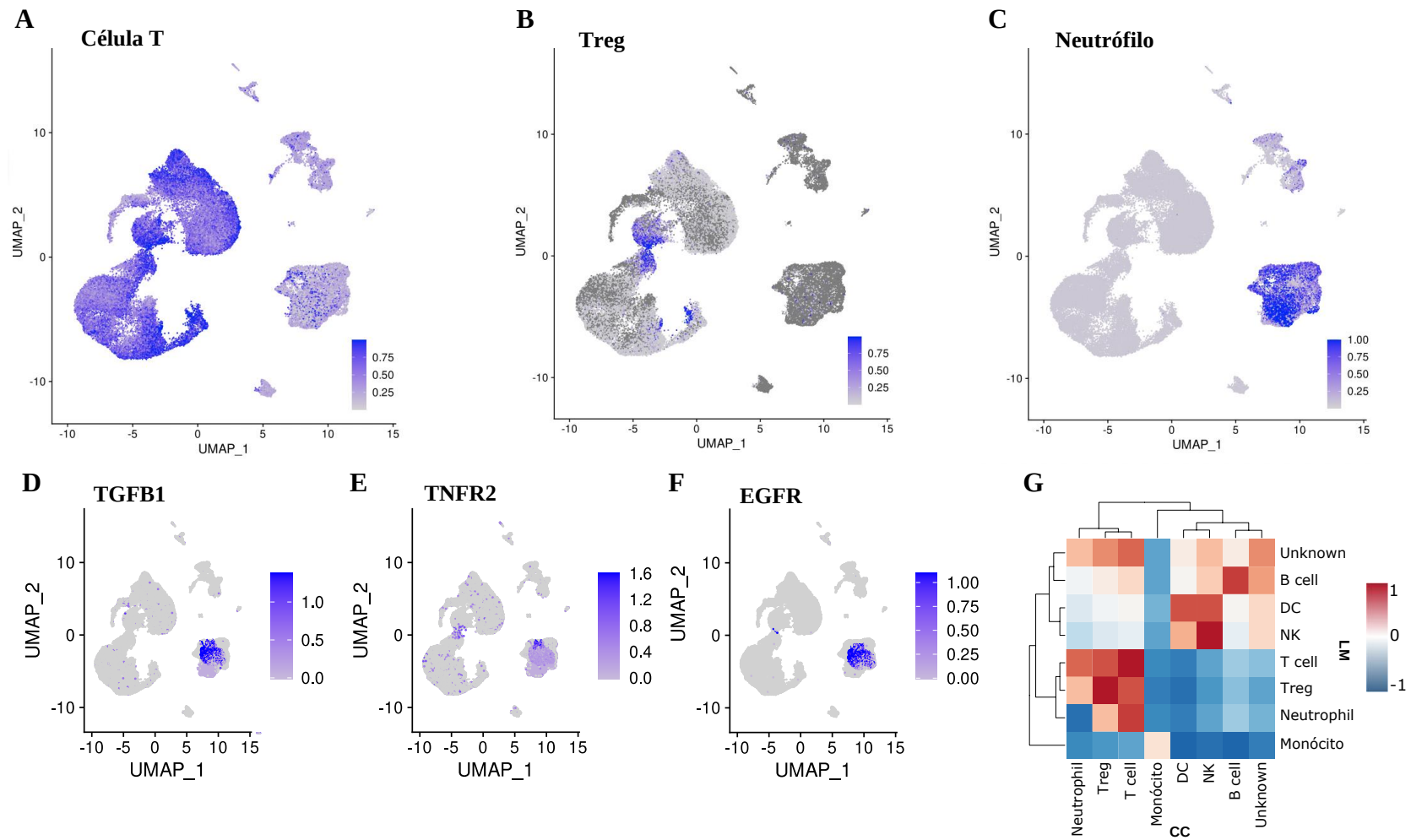


Figura 10 - Representação dos dados de *single-cell RNA-seq* de amostras de metástases hepáticas para a investigação de populações imunes e dos alvos selecionados anteriormente. **(A)** Marcadores de células T mostrando grande presença por toda a metástase e polarização na região periférica dos agrupamentos. **(B)** Representação das células Tregs em regiões específicas do tumor. **(C)** Células FCGR3B+ representando populações de neutrófilos. **(D)** Expressão do gene *TGFB1* com polarização em região marcada pela presença de neutrófilos. **(E)** Expressão do gene *TNFR2* também em agrupamento associado à presença de neutrófilos. Foi observado também a expressão do gene em região com grande presença de células Tregs. **(F)** Marcação de *EGFR* em células agrupadas em região enriquecida para neutrófilos e Tregs. **(G)** Correlação entre populações imunes em LM e CC do mesmo paciente. Observa-se que há correlação negativa para populações de neutrófilos entre os dois conjuntos de dados. CC: câncer primário do cólon; LM: metástases hepáticas ressecáveis.

De todas as populações imunes que foram investigadas, células T (**Figura 10A**), Tregs (**Figura 10B**) e neutrófilos (**Figura 10C**) foram aquelas que apresentaram marcação em regiões próximas, ou sobrepostas, a marcação de 3 dos 12 alvos previamente selecionados, *TGFB1* (**Figura 10D**), *TNFR2* (**Figura 10E**) e *EGFR* (**Figura 10F**). A expressão de *TGFB1* foi maior em populações de células que também expressavam marcadores para neutrófilos. Com relação a *TNFR2*, observa-se maior expressão em dois conjuntos de células marcadas para neutrófilos e Tregs. Essa relação também foi observada para *EGFR* nas metástases, uma vez que é possível observar maior expressão em regiões enriquecidas para neutrófilos e Tregs, apesar desta última aparecer em menor quantidade em comparação ao *TNFR2*. Dessa maneira, sugere-se uma atuação conjunta dessas três proteínas em importantes vias de sinalização cuja atividade pode estar relacionada a um maior recrutamento de células Tregs e neutrófilos nas metástases.

Após a comparação da regulação das 12 proteínas com expressão de transcritos, e confirmação dos achados, realizamos uma segunda análise de confirmação para investigar a regulação destes alvos em dados proteômicos obtidos por MS (EGAD00010002237). Para as 42 amostras investigadas, verificamos em primeiro lugar uma maior atividade das vias EGFR, TGF- β , TNF- α e MAPK nas metástases hepáticas (N=21) comparadas ao tumor primário pareado (N=21) (**Figura 11A**), conforme já havia sido observado anteriormente para dados de mRNA. O cenário também confirmou inibição das vias *trail* e p53 nas metástases em comparação ao tumor primário (**Figura 11A**). Em seguida, confirmamos a maior expressão em metástases de quatro dos 12 alvos selecionados, *TGFB1*, *TNFR2*, *EGFR* e *MDH2* (**Figura 11B**). Ainda, três dos quatro alvos demonstraram uma diferença significativa nos dados do CPTAC de tumores primários e tecido adjacente não-neoplásico do cólon (**Figura 11C**), reforçando um papel de destaque para *TGFB1*, *TNFR2* e *EGFR*. Considerando esses alvos, verificamos que o aumento de expressão de *TGFB1* ($p=0,01$) e *TNFR2* ($p=0,048$) estava associado com pior prognóstico em uma análise de sobrevida global entre pacientes M1 (**Figura 11D**). Por fim, foi criado um classificador utilizando a meta-análise de dados de *microarray* que incluiu 204 amostras de metástases hepáticas e 1.445 tumores primários do cólon para avaliar o potencial prognóstico de ambos os alvos na separação entre tumores primários e metástases. Após as etapas de treinamento (70% das amostras) e teste (30% das amostras), o desempenho na classificação foi maior utilizando *Support Vector Machine* (AUC=0,95), acompanhado de AUC=0,92 para *Random Forest*, AUC=0,81 para KNN (*K Nearest Neighbor*) e AUC=0,75 para *Decision Tree* (**Figura 11E**).

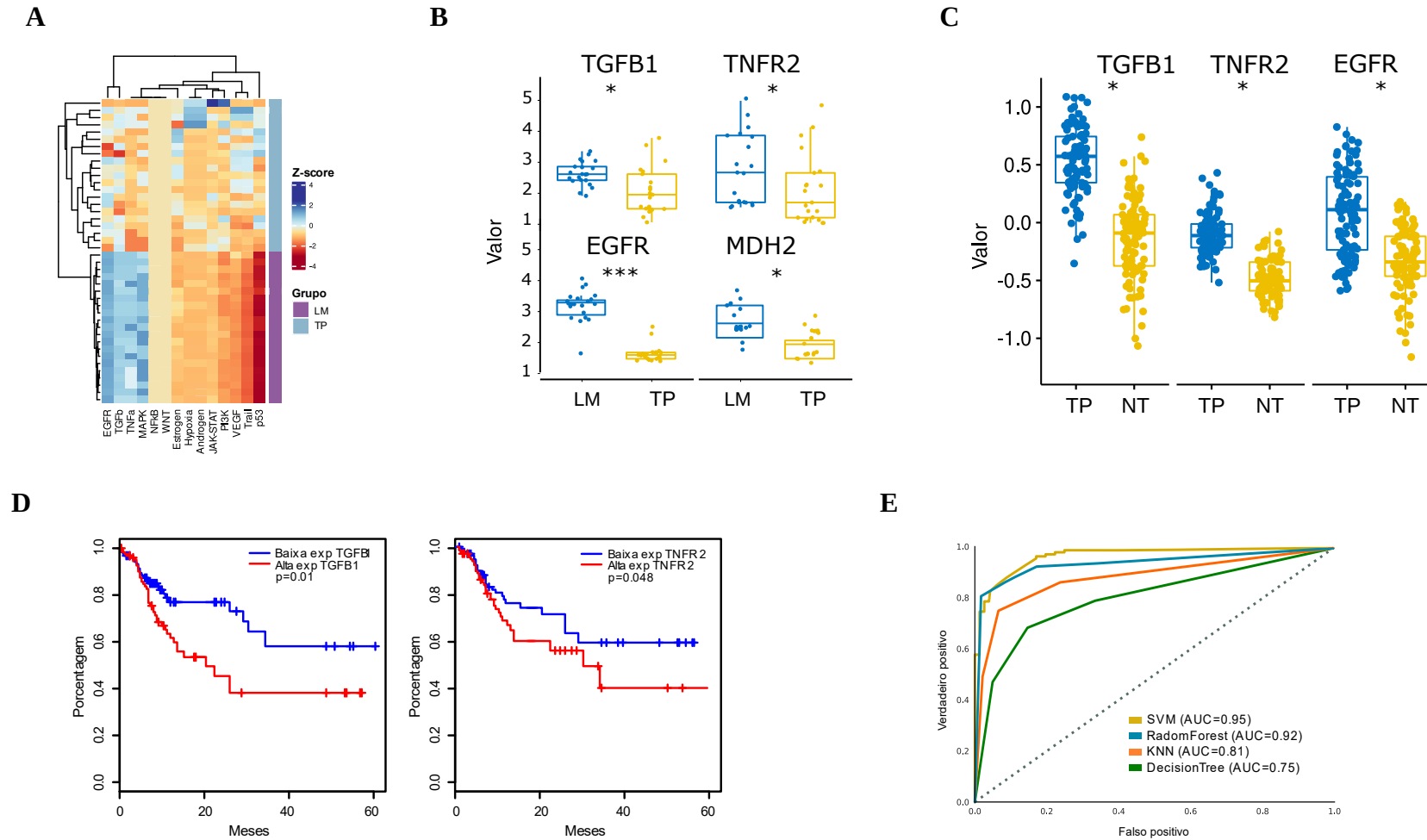


Figura 11 - Confirmação da expressão dos alvos selecionados em conjuntos independentes de amostras. **(A)** No estudo EGAD00010002237, observa-se nas metástases hepáticas forte atividade das vias EGFR, TGF- β , TNF- α e MAPK, assim como inativação de *trail* e p53. **(B)** Além disso, TGFB1, TNFR2, EGFR e MDH2 apresentaram regulação aumentada das proteínas em metástases hepáticas comparadas ao tumor primário. **(C)** Quando considerado o estudo de proteômica (CPTAC), os três primeiros alvos mostraram regulação aumentada significativa em tumores primários em comparação ao tecido não-neoplásico. **(D)** Sobrevida global mostrando a associação da maior expressão de TGFB1 e TNFR2 com pior desfecho clínico. **(E)** Classificador construído utilizando quatro diferentes métodos (*Support Vector Machine* (SVM), *Random Forest*, KNN e *Decision Tree*) e avaliação do desempenho pela AUCROC (*Area Under the ROC curve*). Pelo método SVM foi obtida uma AUC=0,95 para a performance de TGFB1 e TNFR2 na separação de 204 LM e 1.445 TP. TP: câncer primário do cólon; LM: metástases hepáticas ressecáveis; NT: tecido adjacente não-neoplásico.

4.8 PERFIL TRANSCRIPTÔMICO E PROTEÔMICO GLOBAL DAS METÁSTASES RESISTENTES E SENSÍVEIS À TERAPIA

Foi realizada uma comparação específica entre amostras resistentes e sensíveis à terapia sistêmica e cirúrgica, porém, a menor quantidade de dados que passaram pelos critérios de inclusão e exclusão para avaliação de resposta impossibilitou que algumas análises fossem realizadas, como por exemplo, a construção de um classificador com ótimo desempenho. Não foi possível padronizar para uma grande quantidade de amostras o tipo de resposta obtida pelo paciente com o tratamento oferecido. Apesar disso, quatro estudos foram rastreados, sendo eles: GSE3964 (7 metástases sensíveis e 7 resistentes ao tratamento); GSE104645 (30 amostras resistentes e 63 sensíveis ao tratamento que inclui FOLFIRI, FOLFOX, FOLFIRI+bevacizumabe e FOLFOX+bevacizumabe); GSE62322 (7 amostras resistentes e 7 sensíveis ao tratamento); e GSE53127 (12 metástases resistentes e 6 sensíveis ao tratamento). Devido a grande heterogeneidade das informações, cada estudo foi analisado individualmente em comparações entre amostras resistentes e sensíveis. A **Tabela 8** apresenta todos os 57 genes diferencialmente expressos identificados e que foram compartilhados entre todos os estudos. Nesta comparação foi considerado p-valor corrigido para FDR, porém, não foram excluídos os genes de acordo com o critério de $FC \geq 2$ e ≤ -2 . O motivo para a mudança no critério de seleção foi a pouca quantidade de candidatos que seriam mantidos. Dessa maneira, é possível observar a tendência de expressão aumentada ou diminuída para alguns dos possíveis candidatos selecionados para análises posteriores. Assim como na investigação de metástases hepáticas e tumores primários, foi feita uma análise com PCA para cada conjunto de dados buscando por genes com maior variabilidade no conjunto.

Tabela 8 - Genes diferencialmente expressos entre metástases hepáticas do câncer de cólon resistentes e sensíveis ao tratamento.

Gene	GSE3964			GSE104645			GSE62322			GSE53127		
	FC	P-valor	PCA	FC	P-valor	PCA	FC	P-valor	PCA	FC	P-valor	PCA
<i>LHX1</i>	-3,10	2,41E-03	0,45	-2,53	9,12E-08	0,11	-3,00	1,26E-02	0,41	-2,58	1,65E-02	0,11
<i>DBN1</i>	-2,36	1,46E-02	0,87	-2,68	4,17E-06	0,46	-2,57	3,71E-02	0,91	-4,29	4,15E-02	0,47
<i>LAMC1</i>	-2,33	1,35E-03	0,52	-2,32	2,74E-08	0,78	-2,41	8,71E-03	0,48	-2,31	1,10E-02	0,05
<i>PRKCA</i>	-2,27	1,08E-02	0,90	-3,90	2,04E-06	0,22	-4,41	2,99E-02	0,32	-2,16	3,41E-02	0,77
<i>CXCL14#</i>	-2,17	7,32E-03	0,88	-2,65	7,87E-07	0,84	-2,89	2,26E-02	0,90	-5,13	2,68E-02	0,92
<i>SDHA</i>	-2,12	1,84E-02	0,91	-2,24	8,52E-06	0,42	-3,91	4,47E-02	0,91	-5,66	4,88E-02	0,95
<i>KLK1</i>	-2,09	2,83E-04	0,96	-2,75	3,24E-10	0,51	-3,3	1,54E-03	0,63	-4,72	1,25E-03	0,64
<i>DCN</i>	-5,05	1,03E-04	0,35	-3,44	1,84E-12	0,95	-3,17	6,68E-04	0,92	-7,44	1,47E-04	0,95
<i>MDM2</i>	-2,04	1,17E-02	0,72	-2,55	2,48E-06	0,90	-2,95	3,21E-02	0,90	-5,82	3,64E-02	0,98
<i>FAP</i>	-2,09	5,15E-03	0,05	-2,31	2,88E-07	0,78	-2,88	1,81E-02	0,68	-1,30	2,12E-02	0,05
<i>GSK3B</i>	-2,08	7,64E-04	0,76	-2,76	4,20E-09	0,90	-2,79	4,82E-03	0,90	-1,45	5,43E-03	0,92
<i>GM2A</i>	-2,07	9,04E-03	0,45	-2,15	1,32E-06	0,27	-2,7	2,66E-02	0,33	-1,28	3,18E-02	0,81
<i>FLT3LG</i>	-2,04	4,50E-03	0,80	-2,62	2,46E-07	0,12	-2,85	1,68E-02	0,77	-2,97	2,03E-02	0,35
<i>C14ORF9</i>	-2,00	1,28E-02	0,31	-2,42	3,19E-06	0,78	-2,96	3,37E-02	0,39	-1,31	3,90E-02	0,49
<i>EPS15</i>	-2,00	1,95E-02	0,19	-2,19	9,60E-06	0,92	-2,43	4,56E-02	0,57	-3,10	4,97E-02	0,24
<i>DEGS2</i>	-1,91	1,27E-02	0,74	-3,06	3,10E-06	0,54	-2,91	3,37E-02	0,50	-1,53	3,89E-02	0,23
<i>SERPINA1#</i>	-2,77	5,07E-03	0,89	-2,87	2,84E-07	0,86	-2,89	1,80E-02	0,89	-4,82	2,10E-02	0,86
<i>SHB</i>	-1,83	4,60E-03	0,98	-2,74	2,62E-07	0,08	-2,74	1,70E-02	0,84	-3,71	2,03E-02	0,59
<i>ITGAV</i>	-1,88	1,23E-02	0,96	-2,71	2,87E-06	0,77	-2,39	3,29E-02	0,91	-1,10	3,78E-02	0,92
<i>PROM2</i>	-2,79	6,01E-03	0,25	-2,67	4,53E-07	0,52	-2,27	1,97E-02	0,87	-4,85	2,37E-02	0,88
<i>C5</i>	-1,82	7,19E-03	0,87	-2,66	7,03E-07	0,96	-2,51	2,16E-02	0,08	-1,34	2,62E-02	0,42
<i>MED4</i>	-1,97	1,53E-02	0,77	-2,48	4,53E-06	0,67	-3,05	3,80E-02	0,02	-1,13	4,25E-02	0,43
<i>EMX1</i>	-3,86	1,69E-02	0,30	-2,46	6,61E-06	0,86	-2,79	4,19E-02	0,83	-4,44	4,56E-02	0,95
<i>CYP19A1</i>	-1,85	2,98E-04	0,06	-2,46	8,13E-10	0,97	-2,91	2,30E-03	0,98	-3,18	1,74E-03	0,81
<i>CGI-30</i>	-1,99	5,57E-03	0,05	-2,37	4,21E-07	0,12	-2,87	1,90E-02	0,01	-1,78	2,30E-02	0,20
<i>NID</i>	-1,81	7,99E-03	0,52	-2,27	8,90E-07	0,34	-2,65	2,32E-02	0,78	-1,28	2,87E-02	0,15
<i>CCR1</i>	-1,84	1,00E-02	0,78	-2,24	1,70E-06	0,59	-2,61	2,82E-02	0,04	-1,18	3,34E-02	0,44
<i>PECAM1</i>	-1,84	7,69E-03	0,43	-2,22	8,06E-07	0,47	-2,44	2,26E-02	0,57	-7,93	2,70E-02	0,93
<i>FBXW5</i>	-1,79	1,12E-02	0,55	-2,2	2,26E-06	0,72	-2,48	3,13E-02	0,32	-1,05	3,53E-02	0,71
<i>HTR2C</i>	-1,83	2,09E-03	0,17	-2,16	6,65E-08	0,38	-2,62	1,13E-02	0,32	-3,23	1,54E-02	0,85
<i>EIF2S1</i>	-1,72	1,86E-02	0,06	-2,16	8,90E-06	0,57	-2,8	4,49E-02	0,12	-2,00	4,88E-02	0,76
<i>TNFR2#</i>	2,07	5,02E-03	0,90	2,72	2,80E-07	0,93	2,73	1,80E-02	0,95	6,60	2,10E-02	0,90

Con/Tabela 8 - Genes diferencialmente expressos entre metástases hepáticas do câncer de cólon resistentes e sensíveis ao tratamento.

Gene	GSE3964			GSE104645			GSE62322			GSE53127		
	FC	P-valor	PCA	FC	P-valor	PCA	FC	P-valor	PCA	FC	P-valor	PCA
<i>ROR2</i>	2,19	6,81E-04	0,52	2,41	2,99E-09	0,90	2,89	4,08E-03	0,94	3,15	4,31E-03	0,91
<i>NFE2L3</i>	2,98	1,09E-02	0,91	2,61	2,12E-06	0,90	2,74	3,03E-02	0,96	5,24	3,42E-02	0,96
<i>LGR5#</i>	2,12	1,17E-03	0,82	2,55	1,29E-08	0,94	2,97	7,43E-03	0,83	5,06	8,58E-03	0,89
<i>PTGDS</i>	2,18	4,12E-03	0,43	2,82	1,95E-07	0,91	2,51	1,63E-02	0,93	4,14	1,97E-02	0,96
<i>COL4A2</i>	2,10	4,24E-03	0,38	2,28	2,16E-07	0,48	2,44	1,63E-02	0,25	3,36	2,00E-02	0,05
<i>SSB3</i>	2,18	9,35E-03	0,58	3,12	1,62E-06	1,00	2,72	2,78E-02	0,81	2,16	3,28E-02	0,81
<i>TXNDC4</i>	2,24	7,99E-03	0,98	2,69	8,83E-07	0,26	2,64	2,32E-02	0,01	2,03	2,87E-02	0,40
<i>RAB40C</i>	2,05	1,18E-02	0,66	3,12	2,57E-06	0,81	3,00	3,23E-02	0,60	3,20	3,70E-02	0,96
<i>NSMAF</i>	2,10	1,30E-02	0,30	2,45	3,24E-06	0,53	2,56	3,44E-02	0,19	2,53	3,98E-02	0,87
<i>ITIH2</i>	2,03	1,07E-02	0,51	2,53	1,88E-06	0,91	2,37	2,97E-02	0,86	2,17	3,38E-02	0,84
<i>TUBB6</i>	2,09	1,96E-03	0,29	2,49	6,16E-08	0,23	2,48	1,10E-02	0,91	3,49	1,44E-02	0,04
<i>THY1</i>	2,00	3,57E-04	0,66	2,74	1,37E-09	0,41	2,67	2,57E-03	0,62	3,16	2,80E-03	0,58
<i>FOXM1</i>	1,98	5,55E-03	0,55	4,37	3,95E-07	0,96	2,57	1,89E-02	0,91	5,78	2,28E-02	0,93
<i>CDH2#</i>	2,29	5,77E-03	0,96	2,20	4,30E-07	0,87	2,79	1,93E-02	0,85	3,03	2,30E-02	0,84
<i>PAQR4</i>	2,08	1,52E-02	0,78	2,66	4,43E-06	0,57	2,66	3,76E-02	0,83	2,83	4,20E-02	0,10
<i>UFC1</i>	1,94	1,30E-02	0,85	2,59	3,20E-06	0,55	2,58	3,42E-02	0,87	2,60	3,95E-02	0,47
<i>OLFML2A</i>	2,46	1,26E-03	0,91	2,44	1,71E-08	0,58	3,00	7,90E-03	0,91	2,51	9,14E-03	0,54
<i>S100A12#</i>	2,10	8,66E-03	0,97	2,84	1,17E-06	0,84	2,50	2,52E-02	0,87	2,20	3,09E-02	0,84
<i>CLCN2</i>	2,37	3,61E-03	0,04	2,70	1,33E-07	0,61	3,10	1,43E-02	0,91	2,01	1,88E-02	0,08
<i>MCF2</i>	2,23	1,77E-02	0,10	2,62	8,15E-06	0,59	2,43	4,38E-02	0,04	3,89	4,82E-02	0,26
<i>TH</i>	2,04	1,34E-02	0,16	2,60	3,47E-06	0,60	2,69	3,52E-02	0,08	2,57	4,01E-02	0,45
<i>BMP6</i>	2,23	2,85E-03	0,23	2,45	1,01E-07	0,67	2,40	1,33E-02	0,12	2,08	1,70E-02	0,55
<i>PTN</i>	2,17	1,69E-02	0,29	2,45	6,46E-06	0,87	2,49	4,19E-02	0,16	3,46	4,52E-02	0,14
<i>EIF4EBP1</i>	2,04	8,50E-03	0,50	2,63	1,12E-06	0,97	2,78	2,51E-02	0,93	2,22	3,08E-02	0,94

GSE3964: FOLFIRI; GSE104645: FOLFIRI, FOLFOX, FOLFIRI+bevacizumabe ou FOLFOX+bevacizumabe; GSE62322: FOLFIRI; GSE53127: FOLFIRI+bevacizumabe ou CAPOX+bevacizumabe. FOLFOX: 5-fluorouracil, leucovorina e oxaliplatina; FOLFIRI: 5-fluorouracil, leucovorina e irinotecano; CAPOX: capecitabina e oxaliplatina. FC: fold-change; # representa $PC1 \geq 0,80$ para todas as comparações.

Considerando a atividade de sinalização das vias em cada estudo, identificamos em metástases hepáticas resistentes uma ativação significativa das vias TNF- α , NFkB, hipóxia, WNT, VEGF e TGF- β . Apenas no estudo GSE62322 os valores não atingiram $p < 0,05$, apesar do *z-score* concordar com os outros estudos (**Figura 12**).

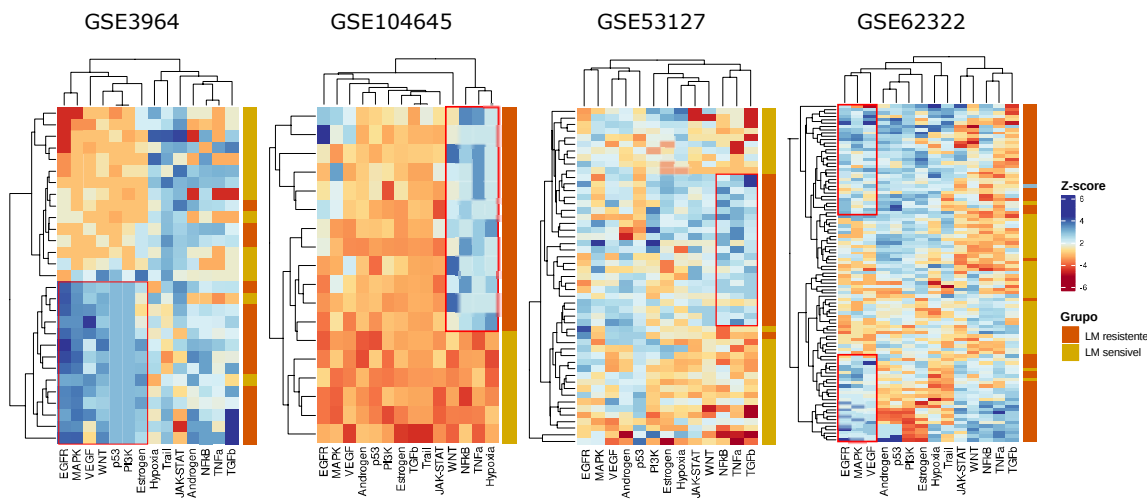


Figura 12 - Atividade das vias de sinalização dos estudos GSE3964, GSE104645, GSE53127 e GSE62322. O *z-score* ≥ 2 define ativação da via (azul) e *z-score* ≤ -2 define inibição (vermelho).

A partir dos 57 genes diferenciais identificados, uma análise de enriquecimento para vias de sinalização do KEGG foi realizada, e aquelas com *p*-valor ajustado para $FDR < 0,05$ foram selecionadas (**Tabela 9**). Assim foram identificadas importantes vias como PI3K/AKT, WNT, ERBB, mTOR e TGF- β

Tabela 9 - Enriquecimento dos genes para vias de sinalização utilizando a base de dados do KEGG.

Vias biológicas	N.genes	adj.Pval	Genes
Via de sinalização PI3K/AKT	8 354	6,03E-06	GSK3B, COL4A2, EIF4EBP1, MDM2, FLT3LG, ITGAV, PRKCA, LAMC1
Via de sinalização WNT	5 166	1,01E-04	GSK3B, TCF7, PRKCA, ROR2, LGR5
Caminhos no câncer	8 531	1,08E-04	GSK3B, COL4A2, TCF7, MDM2, FLT3LG, ITGAV, PRKCA, LAMC1
Adesão focal	5 201	2,46E-04	GSK3B, COL4A2, ITGAV, PRKCA, LAMC1
Complemento e cascatas de coagulação	3 85	1,75E-03	THBD, C5, SERPINA1
Via de sinalização ErbB	3 85	1,75E-03	GSK3B, EIF4EBP1, PRKCA
Interação ECM-receptor	3 88	1,93E-03	COL4A2, ITGAV, LAMC1
Junção de lacuna	3 88	1,93E-03	TUBB6, HTR2C, PRKCA
Câncer de pulmão de pequenas células	3 92	2,19E-03	COL4A2, ITGAV, LAMC1

Con/Tabela 9 - Enriquecimento dos genes para vias de sinalização utilizando a base de dados do KEGG.

Vias biológicas	N.genes	adj.Pval	Genes
Câncer de próstata	3 97	2,54E-03	<i>GSK3B, TCF7, MDM2</i>
Proteoglicanos no câncer	4 205	2,60E-03	<i>MDM2, ITGAV, PRKCA, DCN</i>
Via de sinalização mTOR	3 154	9,21E-03	<i>GSK3B, EIF4EBP1, PRKCA</i>
Câncer colorretal	2 86	2,42E-02	<i>GSK3B, TCF7</i>
Via de sinalização TGF- β	2 94	2,86E-02	<i>DCN, BMP6</i>
Ciclo celular	2 124	4,73E-02	<i>GSK3B, MDM2</i>

Esta análise foi conduzida com os 57 genes identificados entre metástases hepáticas resistentes e sensíveis ao tratamento. Na coluna “N.genes” a barra separa o número de genes alvos identificados em cada módulo e a quantidade de genes que representam toda a via. adj.Pval: p-valor ajustado pela taxa de falso-positivo (FDR, do inglês *False Discovery Rate*).

Com relação ao nosso conjunto de metástases hepáticas irresssecáveis, das 15 amostras inicialmente investigadas apenas 14 foram utilizadas, uma vez que para um paciente não foi possível avaliar o tempo de resposta. Assim, 7 pacientes apresentaram PD em tempo igual ou inferior a 6 meses, sendo considerados resistentes, e outros 7 apresentaram PD entre 8 e 30 meses, sendo considerados sensíveis. Após a comparação entre grupos foram identificadas 32 proteínas com regulação diferencial, considerando p-valor $<0,05$, $FC \geq 2$ e ≤ -2 (**Figura 13**). Apesar do pequeno número de amostras é possível observar a formação de dois grandes agrupamentos, sugerindo essas proteínas como possíveis assinaturas moleculares capazes de identificar metástases com potencial agressivo.

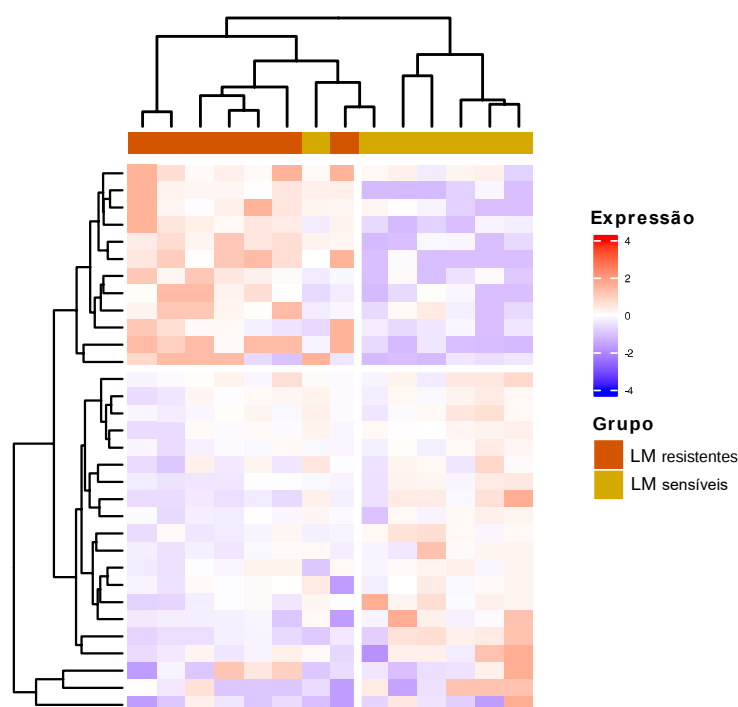


Figura 13 - Identificação de 32 proteínas diferencialmente abundantes após comparação entre os grupos de pacientes que apresentaram PD em tempo igual ou inferior a 6 meses (metástases

resistentes) e pacientes que apresentaram PD após esse período (metástases sensíveis). LM: metástases hepáticas irrecorríveis. Vermelho indica aumento da abundância e azul diminuição.

As proteínas identificadas (18 com abundância aumentada e 14 com abundância diminuída) foram apresentadas na **Tabela 10**. Entre as proteínas com diminuição de abundância nas metástases resistentes o menor FC foi observado para a EMX2 (FC=-4,55, p=4,77E-04). Entre os alvos com aumento de abundância, LGR5 apresentou o maior FC (FC=3,00, p=1,62E-04). Observa-se que SERPINA1, CXCL14, LGR5, CDH2 e S100A12 também foram encontrados como diferencialmente expressos na comparação entre transcritos sensíveis e resistentes ao tratamento.

Tabela 10 - Identificação das 32 proteínas diferencialmente abundantes após comparação entre pacientes que apresentaram PD em tempo igual ou inferior a 6 meses (metástases hepáticas resistentes) e pacientes que apresentaram PD após período de 6 meses (metástases hepáticas sensíveis).

Proteína	FC (resistentes - sensíveis)	adj.Pval
Proteínas com abundância diminuída		
EMX2	-4,55	4,77E-04
PROM1	-4,06	3,66E-04
GPI	-3,58	3,80E-04
TRA2A	-3,10	5,65E-03
RPL23	-3,06	6,56E-03
SDCBP	-2,60	7,72E-02
PFKL	-2,58	1,61E-03
SERPINA1	-2,56	1,70E-03
NDUFS3	-2,53	6,78E-02
TMEM33	-2,52	6,72E-02
IGF2BP3	-2,49	7,78E-03
PTTG1IP	-2,46	3,63E-03
CDX1	-2,40	2,65E-03
CXCL14	-2,33	4,79E-02
Proteínas com abundância aumentada		
LGR5	3,00	1,62E-04
RPL36AL	2,89	1,74E-03
PFKP	2,82	2,66E-03
NHP2L1	2,73	7,70E-03
S100A12	2,71	7,58E-03
PSMB8	2,67	5,71E-03
GPRC5A	2,56	3,75E-02
NONO	2,55	3,64E-02
CDH2	2,55	1,69E-02
C8B	2,54	1,71E-02
RELA	2,51	1,65E-02
TCP1	2,51	2,73E-02
NDUFB10	2,46	3,65E-02
PIR	2,37	3,70E-02
EIF2S1	2,33	4,81E-02
STX7	2,29	4,84E-02
RPL36	2,23	4,62E-02
RPN2	2,20	4,61E-02

FC: fold-change; adj.Pval: p-valor ajustado pela taxa de falso-positivo (FDR, do inglês *False Discovery Rate*).

Em seguida foi realizada uma análise de enriquecimento com os 32 alvos selecionados, separados em proteínas com regulação aumentada e diminuída. Entre aquelas com aumento destacam-se as vias HIF-1 ($p=4,25E-03$), apoptose ($p=7,11E-03$), NFkB ($p=4,86E-02$), IL-17 ($p=4,41E-02$) e sinalização do receptor de células B ($p=4,98E-02$) e T ($p=4,43E-02$). Algumas proteínas aparecem com maior presença na análise, como é o caso de RELA e EIF2S1. Entre as proteínas com diminuição de regulação foi identificada a via de AMPK ($p=4,79E-02$) e também HIF-1 ($p=4,77E-02$) (**Tabela 11**).

Tabela 11 - Enriquecimento para vias biológicas considerando as 32 proteínas diferencialmente abundantes entre metástases sensíveis e resistentes ao tratamento sistêmico.

Vias biológicas	N.proteínas	adj.Pval	Proteínas
Proteínas com abundância aumentada			
Doença do coronavírus	4 232	4,75E-05	RPL36AL, RPL36, C8B, RELA
Doença de Alzheimer	3 369	4,14E-03	NDUFB10, EIF2S1, RELA
Via de sinalização HIF-1	2 109	4,25E-03	RELA, PFKP
Sarampo	2 139	6,82E-03	EIF2S1, RELA
Apoptose	2 142	7,11E-03	EIF2S1, RELA
Vias de neurodegeneração	3 475	8,33E-03	NDUFB10, EIF2S1, RELA
Hepatite C	2 157	8,63E-03	EIF2S1, RELA
Ribossomo	2 158	8,73E-03	RPL36AL, RPL36
Proteassoma	1 46	4,06E-02	PSMB8
Via de sinalização do receptor de células B	1 49	4,32E-02	RELA
Via de sinalização de IL-17	1 50	4,41E-02	RPN2
Via de sinalização do receptor de células T	1 57	4,43E-02	RELA
Via de sinalização do receptor Toll-like	1 63	4,52E-02	RELA
Via de sinalização do receptor de lectina tipo C	1 65	4,69E-02	RELA
Via de sinalização NF kappa B	1 67	4,86E-02	RELA
Via de sinalização AMPK	1 67	4,88E-02	PFKP
Proteínas com abundância diminuída			
Via das Pentoses Fosfato	2 30	1,96E-04	GPI, PFKL
Glicólise Gliconeogênese	2 67	9,80E-04	GPI, PFKL
Metabolismo da galactose	1 31	2,15E-02	PFKL
Metabolismo da frutose e manose	1 33	2,29E-02	PFKL
Metabolismo do amido e da sacarose	1 36	2,49E-02	GPI
Metabolismo do açúcar amino e do açúcar nucleotídeo	1 48	3,31E-02	GPI
Metabolismo central do carbono no câncer	1 70	3,79E-02	PFKL
Degradação de RNA	1 79	3,89E-02	PFKL
Complemento e cascatas de coagulação	1 85	3,94E-02	SERPINA1
Interação de proteína viral com citocina e receptor de citocina	1 100	4,58E-02	CXCL14
Via de sinalização do glucagon	1 107	4,68E-02	PFKL
Via de sinalização HIF-1	1 109	4,77E-02	PFKL
Via de sinalização AMPK	1 120	4,79E-02	PFKL

Na coluna “N.proteínas” a barra separa o número de alvos identificados em cada módulo e a quantidade de alvos que representam toda a via. adj.Pval: p-valor ajustado pela taxa de falso-positivo (FDR, do inglês *False Discovery Rate*).

4.9 ATIVIDADE DA SINALIZAÇÃO CELULAR E PERFIL IMUNE DE METÁSTASES

Considerando essa separação de acordo com a resposta à terapia, não houve agrupamento completo entre amostras sensíveis e resistentes com relação à atividade de vias de sinalização, porém o retângulo vermelho da **Figura 14** apresenta dois grupos de pacientes resistentes com ativação nas vias EGFR ($z\text{-score}=3,69$, $p=3,78E-05$), MAPK ($z\text{-score}=2,76$, $p=1,36E-03$), PI3K ($z\text{-score}=2,27$, $p=4,75E-02$), TGF- β ($z\text{-score}=2,12$, $p=2,11E-02$), VEGF ($z\text{-score}=2,69$, $p=2,28E-02$) e hipóxia ($z\text{-score}=2,06$, $p=4,68E-02$). Também houve uma tendência de diminuição de atividade para a via p53 ($z\text{-score}=-1,98$, $p=1,13E-01$), porém o resultado não foi significativo.

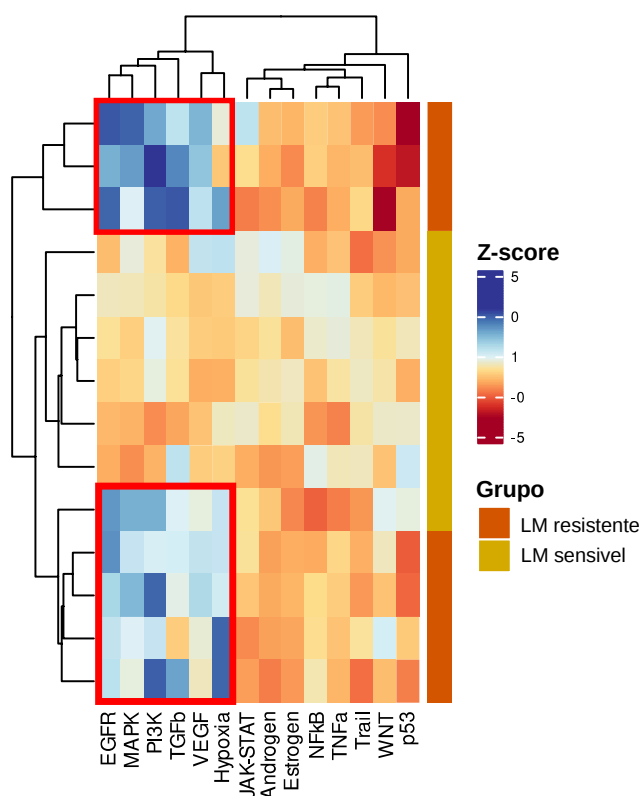


Figura 14 - Heatmap representando a atividade de sinalização entre metástases hepáticas irrissecáveis resistentes e sensíveis à terapia sistêmica. O $z\text{-score} \geq 2$ define atividade aumentada da via (em azul) e $z\text{-score} \leq -2$ atividade diminuída (em vermelho).

A **Tabela 12** apresenta o $z\text{-score}$ referente à atividade das vias em metástases resistentes e sensíveis. Apesar de muitas não serem significativas, é possível notar uma tendência de

diminuição de atividade para estrógeno, JAK-STAT, NFkB, TNF- α , *trail* e WNT em pacientes com perfil de resistência.

Tabela 12 - Perfil de ativação e inibição de vias em metástases sensíveis e resistentes ao tratamento.

Resistentes vs. Sensíveis		
Via	Score	P-valor
Andrógeno	1,91	1,90E-02
EGFR	3,69	3,78E-05
Estrógeno	-1,21	1,06E-02
Hypoxia	2,06	4,68E-02
JAK-STAT	-0,175	7,75E-01
MAPK	2,76	1,36E-03
NFkB	-0,750	1,92E-01
p53	-1,98	1,13E-01
PI3K	2,27	4,75E-02
TGF- β	2,12	2,11E-02
TNF- α	-0,777	1,00E-01
<i>Trail</i>	-1,04	3,16E-02
VEGF	2,69	2,28E-02
WNT	-1,24	1,67E-01

Todos os valores de ativação/inibição de vias foram representados em *z-score*. Dessa maneira, *z-score* ≥ 2 indica ativação e *z-score* ≤ -2 indica inativação.

4.10 PERFIL IMUNOLÓGICO DAS AMOSTRAS INVESTIGADAS

Com relação às populações imunes nas metástases foi observada presença significativa de células Tregs ($p=2,13E-07$) e diminuição de células B ($p=3,44E-04$) em amostras resistentes. Há uma maior tendência de macrófagos M2 em pacientes resistentes, porém esse resultado não foi significativo ($p=1,22E-01$) (**Figura 15**). Por outro lado, pacientes considerados sensíveis apresentaram maior presença de células T CD8+ ($p=3,93E-03$). Nenhuma outra população imune apresentou diferença significativa entre os grupos de metástases.

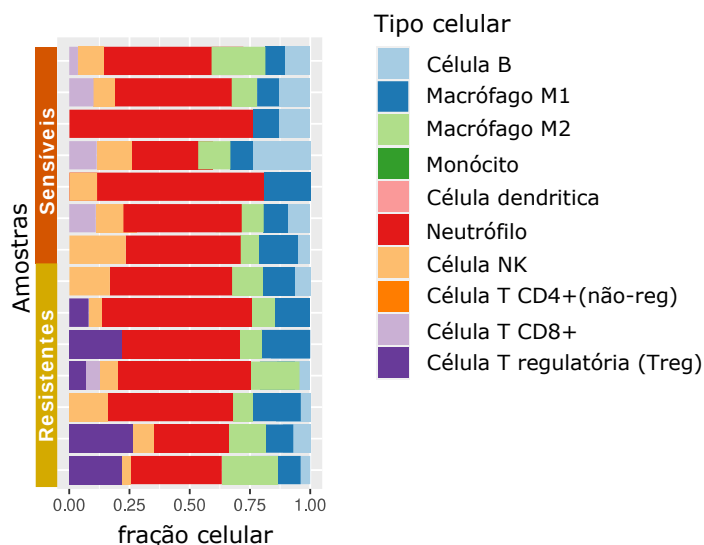


Figura 15 - Perfil imune obtido com as proteínas selecionadas após comparação entre metástases hepáticas irresssecáveis resistentes e sensíveis à terapia sistêmica.

4.11 COMPARAÇÃO DOS RESULTADOS EM CONJUNTOS INDEPENDENTES DE METÁSTASES HEPÁTICAS

Posteriormente, selecionamos um conjunto independente de 63 amostras de metástases hepáticas (GSE81423) obtidas de pacientes com e sem recidiva para comparar o perfil de expressão dos principais alvos observados nas análises anteriores. Dessas, LGR5, CDH2, CDX1, CXCL14 e SERPINA1 tiveram o mesmo perfil de expressão nas metástases que recidivaram, sugerindo uma participação em processos biológicos e vias de sinalização que promovam a resistência das células tumorais aos efeitos da terapia (**Figura 16**).

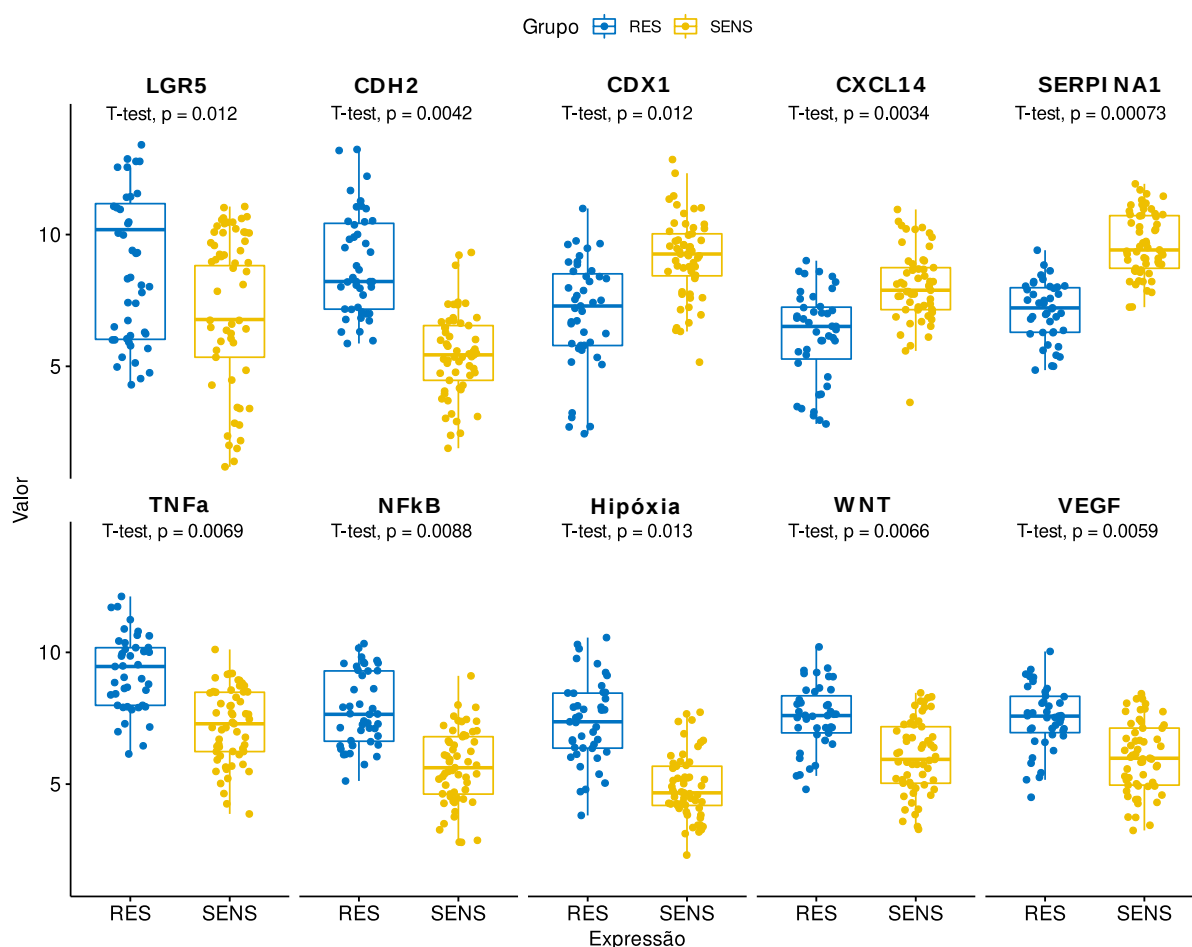


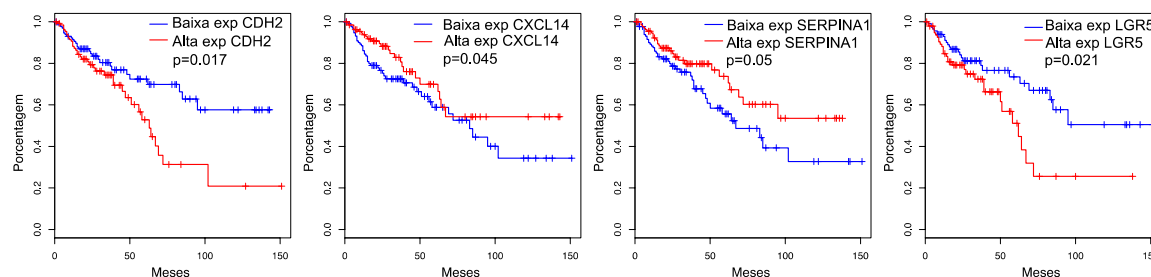
Figura 16 - Comparação de proteínas diferencialmente abundantes entre metástases irrecorríveis resistentes e sensíveis ao tratamento com 63 metástases colorretais com e sem recidiva após terapia obtidas no estudo GSE81423. As proteínas (LGR5, CDH2, CDX1, CXCL14 e SERPINA1) e vias de sinalização (TNF- α , NF κ B, hipóxia, WNT e VEGF) apresentadas foram aquelas que apresentaram o melhor resultado com base no $p < 0,05$.

4.12 IMPACTO DO PERFIL DE EXPRESSÃO DOS ALVOS NA SOBREVIVÊNCIA DOS PACIENTES

As mesmas proteínas que prosseguiram nas análises, LGR5, CDH2, CDX1, CXCL14 e SERPINA1, foram submetidas à análise de sobrevivência global em dois conjuntos de pacientes analisados em relação à expressão gênica. Primeiro, buscou-se uma possível associação prognóstica em uma casuística com 53 amostras de tumores primários M1 disponíveis no TCGA. Nessa análise, nós identificamos *CDH2* ($p = 1,7E-02$), *CXCL14* ($p = 4,5E-02$), *SERPINA1* ($p = 5,0E-02$) e *LGR5* ($p = 2,1E-02$) associados a pior prognóstico nos pacientes (**Figura 17A**). Na segunda etapa, utilizando a mesma abordagem em uma casuística com 127 amostras de metástases hepáticas do estudo GSE159216, os alvos CDH2 ($p = 8,0E-03$), CXCL14

($p=5,0E-02$), SERPINA1 ($p=1,2E-02$) e LGR5 ($p=4,9E-02$) tiveram o perfil de expressão associado significativamente a um desfecho clínico desfavorável (**Figura 17B**).

A



B

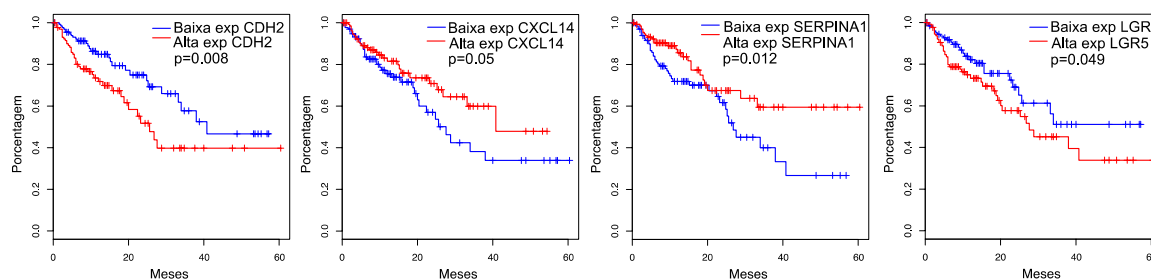


Figura 17 - Sobrevida global de acordo com a expressão dos alvos CDH2, CXCL14, SERPINA1 e LGR5. **(A)** Sobrevida global em tumores primários M1 obtidos no banco de dados do TCGA. **(B)** Sobrevida global em metástases hepáticas de tumores de cólon obtidas do estudo GSE159216. M1: tumores primários metastáticos.

5 DISCUSSÃO

A sobrevida global de 5 anos de pacientes com metástases colorretais submetidos à ressecção hepática aumentou de 20-30% para 40-50% desde as primeiras publicações na década de 1970 (Wilson et al. 1976). Isso provavelmente se deve às melhorias no manejo cirúrgico e à introdução da quimioterapia moderna com uma melhor compreensão da seleção de pacientes. Enquanto a sobrevida a longo prazo foi observada em um grupo de pacientes apresentando metástases múltiplas ou grandes, a recorrência irressuscitável agressiva foi observada em vários pacientes com pequenas metástases solitárias. Dessa maneira, há uma grande oportunidade para elevar o conhecimento da biologia desses tumores com a identificação de novos alvos moleculares que possam auxiliar na identificação de pacientes respondedores e/ou sugerir novas opções terapêuticas. Nesse sentido, este projeto foi desenvolvido e contribuiu com a identificação de alguns importantes candidatos associados a um pior prognóstico nesses pacientes.

Em buscas na literatura é possível encontrar iniciativas interessantes no manejo de grandes conjuntos de dados e investigação do CCR de maneira não isolada. Em 2014, Chang et al. analisaram 12 estudos para coleta de dados do CCR e desenvolveram um modelo preditivo para detecção da doença. Em outra abordagem, Shangkuan e colaboradores (2017) foram responsáveis por avaliar onze estudos e desenvolver uma análise de risco para a identificação de uma assinatura prognóstica com dez genes. Metástases colorretais no fígado também foram exploradas por Liu et al. (2021) durante a investigação de quatro estudos com dados de expressão gênica em busca de genes centrais em associações por redes de interação. Essas abordagens reforçam a robustez da estratégia utilizada neste projeto uma vez que, para o nosso conhecimento, não há na literatura uma meta-análise tão ampla para a investigação de metástases hepáticas do câncer de cólon.

Especificamente sobre a biologia da doença, múltiplas comparações realizadas no estudo destacaram uma possível associação mecanística entre TNFR2, TGFB1 e EGFR com populações de células T, Tregs e neutrófilos. Para TNFR2 e TGFB1, observamos a associação estatística com pior prognóstico em pacientes M1. Esses achados estão intrinsecamente associados a maior atividade de vias como EGFR, WNT, MAPK, NFkB, VEGFR e PI3K e participação de TNFR2 e TGFB1 como moduladores em processos favoráveis ao tumor. Ao mesmo tempo, metástases resistentes ao tratamento compartilharam maior atividade de EGFR,

MAPK, PI3K, e também destacaram um possível envolvimento de TGF- β e hipóxia durante a progressão da doença. Após o aprofundamento das análises em pacientes que apresentaram PD em período ≤ 6 meses, também sugerimos CDH2, CXCL14, SERPINA1 e LGR5 como possíveis fatores prognósticos de pior evolução.

As células imunes especializadas chamadas linfócitos T reguladores (Tregs) se encontram ativas no sistema circulatório e também em nível tecidual, sendo responsáveis pela homeostase imunológica e supressão de ativações excessivas e constitutivas para prevenir respostas autoimunes (Huppert et al. 2022). Após o desenvolvimento de neoantígenos, as Tregs, por apresentarem potente atividade imunossupressora, são reguladas positivamente para conter a atividade citotóxica de outras células imunes (Huppert et al. 2022). Considerando o microambiente hepático, as Tregs são responsáveis tanto por secretar quanto internalizar TGF- β , que se encontra em abundância no tecido (Huppert et al. 2022). As metástases hepáticas resistentes apresentaram maior infiltração dessas células comparadas com as sensíveis, além de ativação da via de sinalização TGF- β em comparação ao tumor primário, que está de acordo com a atividade conjunta reportada anteriormente. Associado a esse fato, a depleção dessas células já foi associada a uma melhora na sobrevida de pacientes e a alta infiltração intratumoral de Tregs está negativamente correlacionada à resposta terapêutica (Sainson et al. 2020).

Considerando o processo envolvido entre a aquisição de fenótipos específicos e o final estabelecimento de células primárias no fígado, células imunes inflamatórias desempenham um importante papel na indução do EMT e, conseqüentemente, de habilidades migratórias no tumor primário através da produção de citocinas como o fator de necrose tumoral alfa (TNF- α , do inglês tumor necrosis factor- α) e fator de transformação do crescimento beta (TGF- β , do inglês transforming growth factor- β) (Ham et al. 2016). Avaliando a atividade de vias de sinalização associadas com a progressão tumoral e envolvidas em mecanismos de resistência, ambas as vias de sinalização apresentaram ativação em tecidos metastáticos comparados aos tecidos adjacentes não-neoplásicos do fígado. A sinalização de TGF- β , que contém mais de 30 componentes envolvidos em sua via, já foi previamente destacada com um papel principal na formação dos tumores secundários, tanto diretamente nas células tumorais quando em diferentes aspectos do nicho neoplásico (Itatani et al. 2019, Marvin et al. 2020). No contexto oncológico, TGF- β pode apresentar um papel duplo, exibindo tanto uma atividade supressora do crescimento tumoral nos estágios iniciais quanto promotora nas fases avançadas da doença (Itatani et al. 2019, Marvin et al. 2020). A ativação constitutiva da sinalização e o aumento de TGF- β em níveis plasmáticos já foi correlacionado com um comportamento biológico agressivo, sendo reportado recidiva do CCR no fígado após ressecção cirúrgica e baixa

sobrevida (Villalba et al. 2017, Marvin et al. 2020). Apesar da via não ser considerada uma boa oportunidade para o desenvolvimento de terapias considerando sua dupla atividade, estudos mostram uma redução da capacidade metastática quando utilizada como alvo de drogas (Villalba et al. 2017). O aumento de expressão e atividade desses alvos nas metástases hepáticas, somado ao fato de que estudos clínicos estão sendo conduzidos visando a sinalização como alvo terapêutico, reforçam a relevância da via no contexto metastático e no prognóstico dos pacientes. Além disso, sua correlação com o aumento de Tregs infiltrados em células resistentes destaca o papel do TME na agressividade e resposta ao tratamento desses pacientes com câncer de cólon.

TNF- α pertence à superfamília de proteínas transmembrana TNF (TNFSF, do inglês tumor necrosis factor superfamily), que apresenta 19 ligantes relacionados estruturalmente que podem interagir a um ou mais dos 29 receptores da TNFSF, regulando processos associados à inflamação, imunidade e sobrevivência celular (Ham et al. 2016). Assim como TGF- β , a via TNF- α também pode apresentar um papel duplo dependente do contexto, tipo e estágio tumoral, com evidências mostrando tanto um efeito favorável ao prognóstico do paciente, facilitando a entrega de drogas e aumentando sua concentração no tecido neoplásico (Ham et al. 2016), quanto um envolvimento com células imunes e citocinas que promovem proliferação, migração e estabelecimento de metástases no fígado (Al Obeed et al. 2014; Ham et al. 2016; Jiao et al. 2014). Em relação ao CCR, essa via já foi fortemente associada com progressão tumoral e recidivas em pacientes M1, sendo que a citocina TNF- α é regulada positivamente (em níveis plasmáticos e teciduais) em estágios avançados em relação aos iniciais (Alotaibi et al. 2021, Yang et al. 2015). Além disso, a ligação de TNF- α aos seus respectivos receptores, TNFR1 and TNFR2, pode ativar outras vias de sinalização como JNK, NF κ B e a cascata de caspases (Alotaibi et al. 2021; Zhao et al. 2018). Essa ativação ainda pode estimular constitutivamente outras vias de sinalização por meio do *crosstalk*, como a interação de JNK com a via MAPK (Alotaibi et al. 2021), que representa um papel importante na resistência aos anti-BRAF. Apesar de já ressaltado na literatura, o envolvimento dessas e de outras vias de sinalização, associado à presença de subpopulações de células imunes no estroma e no tecido tumoral, reforça a participação dessas cascatas em processos relacionados à tumorigênese e, principalmente, ao estabelecimento de metástases hepáticas, enfatizando também o impacto que vias secundárias exercem em ativações não-canônicas de vias já associadas à resistência terapêutica.

Considerando as principais proteínas que apresentaram abundância diminuída em metástases hepáticas resistentes, o fator de transcrição EMX2 e a glicoproteína transmembrana PROM1 (CD133) se destacam em relação às evidências científicas, com indicações de que a

influência desses alvos na resposta terapêutica é generalizada entre diferentes tipos de tumores (Wang et al. 2019). Em um estudo *in vitro* analisando o gene *EMX2*, seu silenciamento aumentou significativamente a resistência de linhagens celulares de câncer de pulmão a três tratamentos quimioterápicos (Yue et al. 2015). Além disso, os autores também confirmaram uma associação do fator de transcrição com características de EMT (Yue et al. 2015). Outros autores confirmaram que, quando em maiores níveis de expressão, o *EMX2* sensibiliza células de câncer de esôfago à cisplatina (Wang et al. 2019). Ainda, o aumento da expressão de *EMX2* em cultura de células primárias de sarcoma mostrou uma redução nas características de células-tronco, com uma inibição da via canônica WNT que foi confirmada com dados *in silico* (Jimenez-García et al. 2021). Em relação ao CCR, o gene apresentou menores níveis de expressão nas metástases hepáticas em comparação aos tumores primários, e esses também apresentam regulação negativa quando relacionados aos seus respectivos tecidos normais pareados (Aykut et al. 2017). Além disso, uma associação significativa foi encontrada entre a baixa expressão de *EMX2* e a ocorrência de metástases hepáticas em pacientes estadio III, com uma diminuição tanto do tempo de sobrevida global quanto livre de doença quando comparado àqueles com maior expressão do gene pareados (Aykut et al. 2017). Esses dados, além de apontarem uma associação positiva entre a expressão e abundância diminuída de *EMX2* e um pior prognóstico, corroboram com o achado dessa proteína no conjunto de metástases resistentes ao tratamento.

Com relação ao PROM1, a proteína codificada pertencente à família das promininas é considerada um importante marcador de células-tronco tumorais (CSCs, do inglês *cancer stem cells*). Dentre os mecanismos de resistência à múltiplas drogas, acredita-se que as CSCs desempenhem um importante papel, principalmente por suas capacidades de autorrenovação e proliferação (Yuan et al. 2020), e diversos estudos nas últimas décadas reportaram o impacto que células CD133+ desempenham no benefício clínico de pacientes em uso de medicamentos como paclitaxel (Li et al. 2021), 5-FU (Zhu et al. 2014) e cisplatina (Li et al. 2021; Lu et al. 2019; Yuan et al. 2020). De forma contraditória aos dados obtidos e apresentados considerando metástases hepáticas resistentes, os achados na literatura apontam que a insensibilidade ao tratamento é associada à alta expressão e, consequentemente, maior abundância de PROM1. Curiosamente, a adesão de células do tumor primário à ECM dos hepatócitos causa uma regulação positiva de genes associados à sobrevivência e proliferação, principalmente em alvos da família EGF e marcadores de CSCs, como PROM1 e LGR5 (Brodt 2016). Assim, a diminuição da expressão do marcador CD133 nessas células metastáticas não-respondedoras ainda é um achado pouco estudado.

Recentemente, Liu et al. (2018) investigaram amostras de tecido não-neoplásico do cólon, tumor primário e metástases hepáticas obtidas de dois estudos depositados na plataforma GEO, identificando genes diferencialmente expressos exclusivos nas metástases. Desses alvos destacados, os autores atribuíram um papel relevante ao gene *SERPINA1* quando avaliado a rede de interação gênica dos genes diferencialmente expressos selecionados por co-expressão. Esse gene, com expressão aumentada nas metástases hepáticas comparadas ao tecido saudável do cólon, foi notado com expressão diminuída em tumores primários provindos do TCGA e associado a um pior prognóstico desses casos analisados. Outro estudo mais recente notou uma relação significativa inversa entre a expressão de *SERPINA1* e dados de sobrevida e estadiamento – por exemplo, quanto mais avançado o estágio avaliado, menor o nível de expressão do gene (Fu et al. 2021). Ainda, avaliando uma coorte de pacientes com câncer de ovário sensíveis e resistentes ao tratamento com platina, o gene foi um dos que apresentou maior diferença na comparação entre os dois grupos, com um aumento de expressão naqueles pacientes não-respondedores ao tratamento (Wu et al. 2016). Considerando o aumento da abundância dessa proteína em metástases hepáticas comparadas ao tecido adjacente não-neoplásico e a regulação negativa da mesma quando considerado os perfis de resposta ao tratamento, com uma abundância diminuída em pacientes resistentes à terapia sistêmica, os dados controversos levantados na literatura atribuem um papel ainda não completamente compreendido da regulação gênica e proteica da *SERPINA1* em metástases hepáticas resistentes.

CDH2, também conhecida como Caderina-2, é um membro da família de moléculas de adesão dependentes de cálcio que medeiam diretamente a adesão célula-célula homotípica e heterotípica. A Caderina-2 é uma caderina tipo I clássica que consiste em 5 domínios extracelulares ligados a um domínio intracelular funcional e já foi descrita pelo seu envolvimento em outros tipos de câncer, incluindo glioma (Chen et al. 2018), onde pacientes com baixa expressão de *CDH2* tiveram um melhor prognóstico e se beneficiaram da terapia com temozolomida. No câncer de próstata metastático ou resistente a castração, uma proporção substancial de pacientes torna-se resistente à enzalutamida após um período de tratamento. Estudos de Lu et al. (2022) mostraram por *small interfering RNA* que as células desses tumores geralmente exibem capacidades proliferativas e migratórias aumentadas, nas quais a Caderina-2 parece desempenhar um papel importante. Em nosso estudo, o aumento de expressão dessa proteína foi associado a um pior prognóstico, o que sugere uma estratégia mais aprofundada para investigar tanto a participação do alvo no desenvolvimento da doença como fator prognóstico que pode ser detectável e auxiliar na terapia dos pacientes.

O ligante 14 da quimiocina CXC (CXCL14), uma quimiocina do tipo CXC conhecida como quimiocina expressa na mama e no rim (BRAK), já foi associada com a ativação de células B e monócitos, promovendo o desenvolvimento tumoral. Efeitos antitumorais do aumento de expressão de CXCL14 já foram observados em estudos sobre câncer de cabeça e pescoço (Ozawa et al. 2006), câncer de mama (Gu et al. 2012), câncer de pulmão (Tessema et al. 2010) e carcinoma hepatocelular (Wang et al. 2013). Kondo et al. (2016) demonstraram que a expressão de *CXCL14* é significativamente diminuída pela ativação da sinalização de EGFR, e que a restauração da expressão de *CXCL14* contribui para o efeito supressor de gefitinib, um inibidor seletivo de RTKs da sinalização de EGFR. Uma vez que identificamos expressão diminuída do alvo nas metástases hepáticas resistentes e associação com pior prognóstico, esses achados sugerem um papel de destaque para a proteína, enfatizando a necessidade de estudos aprofundados sobre sua associação específica com os quimioterápicos/mAbs administrados aos pacientes uma vez que outro membro da família, CXCL13, já foi descrito por Zhang et al. (2019) como mediador da resistência de 5-FU no CCR.

Outra proteína destacada nos dados apresentados é a LGR5, um alvo da via WNT e importante marcador de CSCs provenientes do cólon (Samaržija 2021). Uma causa bem estabelecida e aceita de recidiva do CCR é a falha na erradicação das subpopulações de CSCs, que representam entre 1-10% das células tumorais, visto que essas reestabelecem o tumor após o regime quimioterápico com características moleculares que conferem vantagem considerando as próximas intervenções (Gonzalez-Villarreal et al. 2020; Morgan et al. 2018). Além de LGR5 estar associado a um pior desfecho clínico, estudos apontam que células tumorais apresentam maiores níveis de expressão do que tecidos adjacentes, destacando ainda um aumento em metástases distantes (Morgan et al. 2018). Pacientes que apresentam baixa expressão de LGR5 mostram melhores respostas às terapias baseadas em 5-FU e também maior tempo de sobrevida livre de progressão (Hsu et al. 2013; Morgan et al. 2018; Stanisavljević et al. 2016).

VEGFA, que apresentou menor abundância em metástases sensíveis em comparação com aquelas resistentes ao tratamento, é uma citocina secretada por células tumorais e do estroma para estimular a invasão de células endoteliais e neovascularização. É comumente encontrada em níveis de abundância elevados em diversos tipos de tumores, sendo associada a um pior prognóstico e desfecho clínico por ocorrência de metástase (Kim et al. 2017; Li et al. 2020; Yeh 2019). Li et al. mostraram através de estudos *in vitro* que células de câncer de ovário resistentes apresentavam um aumento de expressão de VEGFA quando comparadas às células não-resistentes. Somado a isso, o tratamento com bevacizumabe em modelo *in vivo* de câncer de pulmão mostrou a aquisição de um perfil de resistência pela regulação positiva de outros

ligantes alternativos, como FGF2 e seus receptores em células estromais (Montemagno e Pagès 2020). Ainda, a hipóxia causada pela inibição da angiogênese também pode regular positivamente a expressão dessas diferentes proteínas envolvidas no processo, o que contribui para a agressividade e recidiva da doença (Ormseth et al. 2022). Dessa forma, além desses dados, o fenótipo de CIN e o *splicing* alternativo podem contribuir no aumento de expressão e resultar em diferentes isoformas de VEGFA, respectivamente, podendo facilitar na perda de eficácia farmacológica pelo mAb e promover uma expansão do conjunto de células tumorais com maior capacidade de iniciação e disseminação

Além do desenvolvimento de um perfil de resistência estar associado à presença de CSCs persistentes ou através de mudanças no cenário transcricional das células tumorais, as células associadas ao TME, como subpopulações imunes, também desempenham importante papel na resposta terapêutica (Bui et al. 2021). Os neutrófilos associados ao tumor (TANs, do inglês *tumor-associated neutrophils*) foram estabelecidos nos últimos anos com uma importante função na predição do desfecho clínico, especialmente na progressão e no processo metastático, em diversos tipos de tumores (Bui et al. 2021, Mizuno et al. 2019, Wu et al. 2020). Em análises de tecido tumoral biopsiado, os TANs foram localizados principalmente em regiões internas da neoplasia em condições de hipóxia, ainda dispondo de características angiogênicas com a expressão e secreção de citocinas, genes que codificam quimiocinas e uma ampla gama de fatores de crescimento, como BV8, MMP-9 e VEGFA (Bui et al. 2021; Mizuno et al. 2019; Wu et al. 2020). De acordo com dados clínicos e os resultados apresentados, as metástases resistentes ao tratamento quimioterápico de primeira linha demonstraram maior infiltração de TANs em comparação aos pacientes sensíveis, o que pode ser atribuído a regulação positiva de VEGFA, também demonstrada anteriormente, e a ativação significativa da via associada à hipóxia.

6 CONCLUSÃO

- A meta-análise desenvolvida neste projeto avaliou mais de 3.5 mil amostras de estudos públicos independentes, obtidas pelas técnicas de RNA-seq, *microarray* e MS, e 130 mil células obtidas por dados de *single-cell* RNA-seq, em um total de 35 *terabytes* de dados utilizando uma estratégia de integração de dados que possibilitou a investigação de características clínicas importantes no contexto oncológico. A combinação e processamento desses dados heterogêneos permitiu aumentar a robustez dos métodos computacionais utilizados e obtenção de características intrínsecas das metástases hepáticas do câncer de cólon.
- A análise de MS realizada em 15 metástases hepáticas inicialmente irressecáveis e 15 tecidos não-neoplásicos adjacentes foi responsável pela geração de dados inéditos sobre a biologia da doença e permitiu a identificação de TNFR2, EGFR e TGFB1 como possíveis moduladores da atividade de sinalização nas metástases. Essa possível atividade ainda estaria associada à presença de populações de neutrófilos e células Tregs, seja no recrutamento ou em atividades que ainda precisam ser exploradas, o que poderia futuramente ser avaliado em novas estratégias de tratamento. A presença de TNFR2 e TGFB1 foi, inclusive, identificada como possível marcador prognóstico com 95% de acurácia na separação entre câncer primário de cólon e metástases hepáticas.
- Especificamente sobre as características das metástases inicialmente ressecáveis e irressecáveis, foram identificados alguns importantes alvos como YWHAZ, AHNAK, HNRNPH1, PCBP1 e TGFB1 com diferença maior que 10-*fold* entre os grupos de amostras, porém apenas TGFB1 prosseguiu nas análises para ser explorado como potencial marcador. Destacamos também que análises não apresentadas, utilizadas na etapa exploratória, não identificou características imunes ou na atividade da sinalização que pudessem caracterizar as metástases irressecáveis como uma entidade biologicamente diferente das outras metástases.
- Apesar de termos identificado importantes genes como *LGR5*, *SERPINA1*, *CDX1*, *CXCL14* e *CDH2* associados a mecanismos de resistência ao tratamento, nossa casuística

impossibilitou uma análise precisa dos resultados. Com relação à proteômica das metástases inicialmente irressecáveis, ainda foi possível separar os 15 pacientes em grupos com progressão inferior ou superior a 6 meses para definição daqueles com tumores resistentes e sensíveis, respectivamente. Apesar de identificarmos as cinco proteínas em dados de transcritos e associarmos quatro delas (LGR5, SERPINA1, CDH2 e CXCL14) com pior prognóstico, novas análises precisam ser conduzidas para explorar o potencial dessas informações no contexto clínico-biológico.

7 REFERÊNCIAS

Adam R, Kitano Y, Abdelrafee A, Allard MA, Baba H. Debulking surgery for colorectal liver metastases: Foolish or chance? *Surg Oncol.* 2020, 33:266-269. doi:10.1016/j.suronc.2020.02.008.

Adam R, Kitano Y. Multidisciplinary approach of liver metastases from colorectal cancer. *Ann Gastroenterol Surg.* 2019, 3(1):50-56.

Al Obeed OA, Alkhayal KA, Al Sheikh A, et al. Increased expression of tumor necrosis factor- α is associated with advanced colorectal cancer stages. *World J Gastroenterol.* 2014, 20(48):18390-18396. doi:10.3748/wjg.v20.i48.18390.

Alotaibi AG, Li JV, Gooderham NJ. Tumour necrosis factor- α (TNF- α) enhances dietary carcinogen-induced DNA damage in colorectal cancer epithelial cells through activation of JNK signaling pathway. *Toxicology.* 2021, 457:152806. doi:10.1016/j.tox.2021.152806.

Aykut B, Ochs M, Radhakrishnan P, et al. EMX2 gene expression predicts liver metastasis and survival in colorectal cancer. *BMC Cancer.* 2017, 17(1):555. Published 2017 Aug 22. doi:10.1186/s12885-017-3556-2.

Baek SK. Laterality: Right-Sided and Left-Sided Colon Cancer. *Ann Coloproctol.* 2017, 33(6):205-206. doi:10.3393/ac.2017.33.6.205.

Baran B, Mert Ozupek N, Yerli Tetik N, Acar E, Bekcioglu O, Baskin Y. Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. *Gastroenterology Res.* 2018, 11(4):264-273.

Biller LH, Schrag D. Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review. *JAMA.* 2021, 325(7):669-685. doi:10.1001/jama.2021.0106.

Blank-Landeshammer B, Richard VR, Mitsa G, et al. Proteogenomics of Colorectal Cancer Liver Metastases: Complementing Precision Oncology with Phenotypic Data. *Cancers (Basel).* 2019, 11(12):1907. Published 2019 Dec 1. doi:10.3390/cancers11121907.

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976, 72:248-254. doi:10.1006/abio.1976.9999.

Bray SM, Lee J, Kim ST, et al. Genomic characterization of intrinsic and acquired resistance to cetuximab in colorectal cancer patients. *Sci Rep.* 2019, 9(1):15365. Published 2019 Oct 25. doi:10.1038/s41598-019-51981-5.

Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet.* 2014, 383(9927):1490-1502. doi:10.1016/S0140-6736(13)61649-9.

Brodt P. Role of the Microenvironment in Liver Metastasis: From Pre- to Prometastatic Niches. *Clin Cancer Res.* 2016, 22(24):5971-5982. doi:10.1158/1078-0432.CCR-16-0460.

Bui TM, Yalom LK, Sumagin R. Tumor-associated neutrophils: orchestrating cancer pathobiology and therapeutic resistance. *Expert Opin Ther Targets*. 2021, 25(7):573-583. doi:10.1080/14728222.2021.1954162.

Buikhuisen JY, Torang A, Medema JP. Exploring and modelling colon cancer inter-tumour heterogeneity: opportunities and challenges. *Oncogenesis*. 2020, 9(7):66. Published 2020 Jul 9. doi:10.1038/s41389-020-00250-6.

Calderón-Celis F, Encinar JR, Sanz-Medel A. Standardization approaches in absolute quantitative proteomics with mass spectrometry. *Mass Spectrom Rev*. 2018, 37(6):715-737. doi:10.1002/mas.21542.

Cancer Genome Atlas Network (TCGA). Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012, 487(7407):330-337. Published 2012 Jul 18. doi:10.1038/nature11252.

Caputo F, Santini C, Bardasi C, et al. BRAF-Mutated Colorectal Cancer: Clinical and Molecular Insights. *Int J Mol Sci*. 2019, 20(21):5369. Published 2019 Oct 28. doi:10.3390/ijms20215369.

Chan DKH, Lim T-Z, Tan K-K. T4N0 colon cancers should be treated like T3N1 disease. *J Gastrointest Oncol*. 2018, 10(1):6–11.

Chang YT, Yao CT, Su SL et al. Verification of gene expression profiles for colorectal cancer using 12 internet public microarray datasets. *World J Gastroenterol* 2014; 20(46): 17476-17482. doi: 10.3748/wjg.v20.i46.17476

Chen S, Yue T, Huang Z, et al. Inhibition of hydrogen sulfide synthesis reverses acquired resistance to 5-FU through miR-215-5p-EREG/TYMS axis in colon cancer cells. *Cancer Lett*. 2019, 466:49-60. doi:10.1016/j.canlet.2019.09.006.

Chen Q, Cai J, Jiang C. CDH2 expression is of prognostic significance in glioma and predicts the efficacy of temozolomide therapy in patients with glioblastoma. *Oncol Lett*. 2018 May;15(5):7415-7422. doi: 10.3892/ol.2018.8227.

Chow FC, Chok KS. Colorectal liver metastases: An update on multidisciplinary approach. *World J Hepatol*. 2019, 11(2):150-172. doi:10.4254/wjh.v11.i2.150.

Chu QD, Zhou M, Medeiros K, Peddi P. Positive surgical margins contribute to the survival paradox between patients with stage IIB/C (T4N0) and stage IIIA (T1-2N1, T1N2a) colon cancer. *Surgery*. 2016, 160(5):1333-1343.

Cicero G, Addeo R, De Luca R, et al. TAS-102 in metastatic colorectal cancer (mCRC): efficacy, tolerability, and quality of life in heavily pretreated elderly patients: a real-life study. *Drugs Context*. 2020, 9:2020-6-3. Published 2020 Sep 18. doi:10.7573/dic.2020-6-3.

Coebergh van den Braak RRJ, Ten Hoon S, Sieuwerts AM, Tuynman JB, Smid M, Wilting SM, et al. Interconnectivity between molecular subtypes and tumor stage in colorectal cancer. *BMC Cancer*. 2020, 20(1):850.

Colaprico A, Silva TC, Olsen C, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016, 44(8):e71. doi:10.1093/nar/gkv1507.

Corcoran RB, André T, Atreya CE, et al. Combined BRAF, EGFR, and MEK Inhibition in Patients with *BRAF*^{V600E}-Mutant Colorectal Cancer. *Cancer Discov.* 2018, 8(4):428-443. doi:10.1158/2159-8290.CD-17-1226.

Corcoran RB, Ebi H, Turke AB, et al. EGFR-mediated re-activation of MAPK signaling contributes to insensitivity of BRAF mutant colorectal cancers to RAF inhibition with vemurafenib. *Cancer Discov.* 2012, 2(3):227-235. doi:10.1158/2159-8290.CD-11-0341.

Coscia F, Doll S, Bech JM, et al. A streamlined mass spectrometry-based proteomics workflow for large-scale FFPE tissue analysis. *J Pathol.* 2020, 251(1):100-112. doi:10.1002/path.5420.

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res.* 2011, 10(4):1794-1805. doi:10.1021/pr101065j.

Cremolini C, Antoniotti C, Stein A, et al. Individual Patient Data Meta-Analysis of FOLFOXIRI Plus Bevacizumab Versus Doublets Plus Bevacizumab as Initial Therapy of Unresectable Metastatic Colorectal Cancer [published online ahead of print, 2020 Aug 20]. *J Clin Oncol.* 2020, JCO2001225. doi:10.1200/JCO.20.01225.

Cremolini C, Schirripa M, Antoniotti C, et al. First-line chemotherapy for mCRC—a review and evidence-based algorithm. *Nat Rev Clin Oncol.* 2015, 12(10):607-619. doi:10.1038/nrclinonc.2015.129.

Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet.* 2019, 394(10207):1467-1480. doi:10.1016/S0140-6736(19)32319-0.

Del Carmen S, Sayagués JM, Bengoechea O, et al. Spatio-temporal tumor heterogeneity in metastatic CRC tumors: a mutational-based approach. *Oncotarget.* 2018, 9(76):34279-34288. Published 2018 Sep 28. doi:10.18632/oncotarget.26081.

Edwards NJ, Oberti M, Thangudu RR, et al. The CPTAC Data Portal: A Resource for Cancer Proteomics Research. *J Proteome Res.* 2015, 14(6):2707-2713. doi:10.1021/pr501254j.

Ellis MJ, Gillette M, Carr SA, et al. Connecting genomic alterations to cancer biology with proteomics: the NCI Clinical Proteomic Tumor Analysis Consortium. *Cancer Discov.* 2013, 3(10):1108-1112. doi:10.1158/2159-8290.CD-13-0219.

Fahrner M, Bronsert P, Fichtner-Feigl S, Jud A, Schilling O. Proteome biology of primary colorectal carcinoma and corresponding liver metastases. *Neoplasia.* 2021, 23(12):1240-1251. doi:10.1016/j.neo.2021.10.005.

Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell.* 1990, 61(5):759-767. doi:10.1016/0092-8674(90)90186-i.

Franczak C, Witz A, Geoffroy K, et al. Evaluation of KRAS, NRAS and BRAF mutations detection in plasma using an automated system for patients with metastatic colorectal cancer. *PLoS One*. 2020, 15(1):e0227294. Published 2020 Jan 15. doi:10.1371/journal.pone.0227294.

Fu C, Yu Z, He Y, Ding J, Wei M. Down-Regulation of an Autophagy-Related Gene *SERPINA1* as a Superior Prognosis Biomarker Associates with Relapse and Distant Metastasis in Colon Adenocarcinoma. *Onco Targets Ther*. 2021, 14:3861-3872. Published 2021 Jun 22. doi:10.2147/OTT.S306405.

Gao XH, Yu GY, Gong HF, et al. Differences of protein expression profiles, KRAS and BRAF mutation, and prognosis in right-sided colon, left-sided colon and rectal cancer. *Sci Rep*. 2017, 7(1):7882. Published 2017 Aug 11. doi:10.1038/s41598-017-08413-z.

Gbolahan O, O'Neil B. Update on systemic therapy for colorectal cancer: biologics take sides. *Transl Gastroenterol Hepatol*. 2019, 4:9. Published 2019 Feb 13. doi:10.21037/tgh.2019.01.12.

Ghidini M, Petrelli F, Tomasello G. Right Versus Left Colon Cancer: Resectable and Metastatic Disease. *Curr Treat Options Oncol*. 2018, 19(6):31.

Gonzalez-Villarreal CA, Quiroz-Reyes AG, Islas JF, Garza-Treviño EN. Colorectal Cancer Stem Cells in the Progression to Liver Metastasis. *Front Oncol*. 2020, 10:1511. Published 2020 Aug 20. doi:10.3389/fonc.2020.01511.

Grothey A, Van Cutsem E, Sobrero A, et al. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet*. 2013, 381(9863):303-312. doi:10.1016/S0140-6736(12)61900-X.

GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013, 45(6):580-585. doi:10.1038/ng.2653.

Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016, 32(18):2847-2849.

Gu XL, Ou ZL, Lin FJ et al. Expression of CXCL14 and its anticancer role in breast cancer. *Breast Cancer Res Treat*. 2012 Oct;135(3):725-35. doi: 10.1007/s10549-012-2206-2

Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015, 21(11):1350-1356.

Ham B, Fernandez MC, D'Costa Z, Brodt P. The diverse roles of the TNF axis in cancer progression and metastasis. *Trends Cancer Res*. 2016, 11(1):1-27.

Hartung F, Wang Y, Aronow B, Weber GF. A core program of gene expression characterizes cancer metastases. *Oncotarget*. 2017, 8(60):102161-102175. Published 2017 Nov 2. doi:10.18632/oncotarget.22240.

He L, Zhu H, Zhou S, et al. Wnt pathway is involved in 5-FU drug resistance of colorectal cancer cells. *Exp Mol Med*. 2018, 50(8):1-12. Published 2018 Aug 14. doi:10.1038/s12276-018-0128-8.

Hsu HC, Liu YS, Tseng KC, et al. Overexpression of Lgr5 correlates with resistance to 5-FU-based chemotherapy in colorectal cancer. *Int J Colorectal Dis*. 2013, 28(11):1535-1546. doi:10.1007/s00384-013-1721-x.

Huppert LA, Green MD, Kim L, et al. Tissue-specific Tregs in cancer metastasis: opportunities for precision immunotherapy. *Cell Mol Immunol*. 2022, 19(1):33-45. doi:10.1038/s41423-021-00742-4.

Isella C, Brundu F, Bellomo SE, et al. Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nat Commun*. 2017, 8:15107. Published 2017 May 31. doi:10.1038/ncomms15107.

Itatani Y, Kawada K, Sakai Y. Transforming Growth Factor- β Signaling Pathway in Colorectal Cancer and Its Tumor Microenvironment. *Int J Mol Sci*. 2019, 20(23):5822. Published 2019 Nov 20. doi:10.3390/ijms20235822.

Jeon J, Du M, Schoen RE, et al. Determining Risk of Colorectal Cancer and Starting Age of Screening Based on Lifestyle, Environmental, and Genetic Factors. *Gastroenterology*. 2018, 154(8):2152-2164.e19. doi:10.1053/j.gastro.2018.02.021.

Jiang Y, Yan X, Liu K, Shi Y, Wang C, Hu J, et al. Discovering the molecular differences between right- and left-sided colon cancer using machine learning methods. *BMC Cancer*. 2020, 20(1):1012.

Jiao SF, Sun K, Chen XJ, et al. Inhibition of tumor necrosis factor alpha reduces the outgrowth of hepatic micrometastasis of colorectal tumors in a mouse model of liver ischemia-reperfusion injury. *J Biomed Sci*. 2014, 21(1):1. Published 2014 Jan 8. doi:10.1186/1423-0127-21-1.

Jimenez-García MP, Lucena-Cacace A, Otero-Albiol D, Carnero A. Empty spiracles homeobox genes EMX1 and EMX2 regulate WNT pathway activation in sarcomagenesis. *J Exp Clin Cancer Res*. 2021, 40(1):247. Published 2021 Aug 7. doi:10.1186/s13046-021-02048-9.

Kalyan A, Kircher S, Shah H, Mulcahy M, Benson A. Updates on immunotherapy for colorectal cancer. *J Gastrointest Oncol*. 2018, 9(1):160-169. doi:10.21037/jgo.2018.01.17.

Kamal Y, Schmit SL, Hoehn HJ, Amos CI, Frost HR. Transcriptomic Differences between Primary Colorectal Adenocarcinomas and Distant Metastases Reveal Metastatic Colorectal Cancer Subtypes. *Cancer Res*. 2019, 79(16):4227-4241. doi:10.1158/0008-5472.CAN-18-3945.

Kennedy AS, Jarboui MA, Srihari S, Raso C, Bryan K, Dernayka L, et al. Extensive rewiring of the EGFR network in colorectal cancer cells expressing transforming levels of KRASG13D. *Nat Commun*. 2020, 11(1):499.

Kim EK, Song MJ, Jung Y, Lee WS, Jang HH. Proteomic Analysis of Primary Colon Cancer and Synchronous Solitary Liver Metastasis. *Cancer Genomics Proteomics*. 2019, 16(6):583-592. doi:10.21873/cgp.20161.

Kim M, Jang K, Miller P, et al. VEGFA links self-renewal and metastasis by inducing Sox2 to repress miR-452, driving Slug. *Oncogene*. 2017, 36(36):5199-5211. doi:10.1038/onc.2017.4.

Kondo T, Ozawa S, Ikoma T et al. Expression of the chemokine CXCL14 and cetuximab-dependent tumour suppression in head and neck squamous cell carcinoma. *Oncogenesis*. 2016 Jul 11;5(7):e240. doi: 10.1038/oncsis.2016.43.

Koveitypour Z, Panahi F, Vakilian M, et al. Signaling pathways involved in colorectal cancer progression. *Cell Biosci*. 2019, 9:97. Published 2019 Dec 2. doi:10.1186/s13578-019-0361-4.

Krishnamurthy N, Kurzrock R. Targeting the Wnt/beta-catenin pathway in cancer: Update on effectors and inhibitors. *Cancer Treat Rev*. 2018, 62:50-60. doi:10.1016/j.ctrv.2017.11.002.

Ku X, Xu Y, Cai C, Yang Y, Cui L, Yan W. In-Depth Characterization of Mass Spectrometry-Based Proteomic Profiles Revealed Novel Signature Proteins Associated with Liver Metastatic Colorectal Cancers. *Anal Cell Pathol (Amst)*. 2019, 2019:7653230. Published 2019 Nov 4. doi:10.1155/2019/7653230.

Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. *Nat Rev Dis Primers*. 2015, 1:15065. Published 2015 Nov 5. doi:10.1038/nrdp.2015.65.

Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44(W1):W90-W97. doi:10.1093/nar/gkw377.

Lam VW, Spiro C, Laurence JM, et al. A systematic review of clinical response and survival outcomes of downsizing systemic chemotherapy and rescue liver surgery in patients with initially unresectable colorectal liver metastases. *Ann Surg Oncol*. 2012, 19(4):1292-1301. doi:10.1245/s10434-011-2061-0.

Le DT, Uram JN, Wang H, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med*. 2015, 372(26):2509-2520. doi:10.1056/NEJMoa1500596.

Lee YT, Tan YJ, Oon CE. Molecular targeted therapy: Treating cancer with specificity. *Eur J Pharmacol*. 2018, 834:188-196. doi:10.1016/j.ejphar.2018.07.034.

Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012, 28(6):882-883. doi:10.1093/bioinformatics/bts034.

Li C, Sun YD, Yu GY, et al. Integrated Omics of Metastatic Colorectal Cancer. *Cancer Cell*. 2020, 38(5):734-747.e9. doi:10.1016/j.ccell.2020.08.002.

Li J, Qin S, Xu R, et al. Regorafenib plus best supportive care versus placebo plus best supportive care in Asian patients with previously treated metastatic colorectal cancer

(CONCUR): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol.* 2015, 16(6):619-629. doi:10.1016/S1470-2045(15)70156-7.

Li X, Hu Z, Shi H, Wang C, Lei J, Cheng Y. Inhibition of VEGFA Increases the Sensitivity of Ovarian Cancer Cells to Chemotherapy by Suppressing VEGFA-Mediated Autophagy. *Onco Targets Ther.* 2020, 13:8161-8171. Published 2020 Aug 18. doi:10.2147/OTT.S250392.

Li Y, Wang Z, Ajani JA, Song S. Drug resistance and Cancer stem cells. *Cell Commun Signal.* 2021, 19(1):19. Published 2021 Feb 15. doi:10.1186/s12964-020-00627-5.

Lindhorst PH, Hummon AB. Proteomics of Colorectal Cancer: Tumors, Organoids, and Cell Cultures-A Minireview. *Front Mol Biosci.* 2020, 7:604492. Published 2020 Dec 10. doi:10.3389/fmolb.2020.604492.

Liu J, Cho YB, Hong HK, et al. Molecular dissection of CRC primary tumors and their matched liver metastases reveals critical role of immune microenvironment, EMT and angiogenesis in cancer metastasis. *Sci Rep.* 2020a, 10(1):10725. Published 2020 Jul 1. doi:10.1038/s41598-020-67842-5.

Liu J, Lichtenberg T, Hoadley KA, et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell.* 2018, 173(2):400-416.e11. doi:10.1016/j.cell.2018.02.052.

Liu J, Wang D, Zhang C, et al. Identification of liver metastasis-associated genes in human colon carcinoma by mRNA profiling. *Chin J Cancer Res.* 2018, 30(6):633-646. doi:10.21147/j.issn.1000-9604.2018.06.08.

Liu WQ, Li WL, Ma SM, Liang L, Kou ZY, Yang J. Discovery of core gene families associated with liver metastasis in colorectal cancer and regulatory roles in tumor cell immune infiltration. *Transl Oncol.* 2021, 14(3):101011. doi:10.1016/j.tranon.2021.101011.

Liu X, Xu D, Liu Z, et al. THBS1 facilitates colorectal liver metastasis through enhancing epithelial-mesenchymal transition. *Clin Transl Oncol.* 2020b, 22(10):1730-1740. doi:10.1007/s12094-020-02308-8.

Liu S, Zhang Y, Zhang S et al. Identification of Hub Genes Related to Liver Metastasis of Colorectal Cancer by Integrative Analysis. *Front Oncol.* 2021 Aug 19;11:714866. doi:10.3389/fonc.2021.714866

Lochhead P, Imamura Y, Morikawa T, et al. Insulin-like growth factor 2 messenger RNA binding protein 3 (IGF2BP3) is a marker of unfavourable prognosis in colorectal cancer. *Eur J Cancer.* 2012, 48(18):3405-3413. doi:10.1016/j.ejca.2012.06.021.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014, 15(12):550. doi:10.1186/s13059-014-0550-8.

Lu R, Zhao G, Yang Y, Jiang Z, Cai J, Hu H. Inhibition of CD133 Overcomes Cisplatin Resistance Through Inhibiting PI3K/AKT/mTOR Signaling Pathway and Autophagy in CD133-Positive Gastric Cancer Cells. *Technol Cancer Res Treat.* 2019, 18:1533033819864311. doi:10.1177/1533033819864311.

Lu CH, Wu CH, Hsieh PF. Small interfering RNA targeting N-cadherin regulates cell proliferation and migration in enzalutamide-resistant prostate cancer. *Oncol Lett*. 2022 Mar;23(3):90. doi: 10.3892/ol.2022.13210.

Ma YS, Huang T, Zhong XM, et al. Proteogenomic characterization and comprehensive integrative genomic analysis of human colorectal cancer liver metastasis [published correction appears in *Mol Cancer*. 2019 Apr 2, 18(1):72]. *Mol Cancer*. 2018, 17(1):139. Published 2018 Sep 21. doi:10.1186/s12943-018-0890-1.

Martin J, Petrillo A, Smyth EC, et al. Colorectal liver metastases: Current management and future perspectives. *World J Clin Oncol*. 2020, 11(10):761-808. doi:10.5306/wjco.v11.i10.761.

Marvin DL, Heijboer R, Ten Dijke P, Ritsma L. TGF- β signaling in liver metastasis. *Clin Transl Med*. 2020, 10(7):e160. doi:10.1002/ctm2.160.

Matthiesen R, Bunkenborg J. Introduction to mass spectrometry-based proteomics. *Methods Mol Biol*. 2013, 1007:1-45. doi:10.1007/978-1-62703-392-3_1.

McInnes T, Zou D, Rao DS, et al. Genome-wide methylation analysis identifies a core set of hypermethylated genes in CIMP-H colorectal cancer. *BMC Cancer*. 2017, 17(1):228. Published 2017 Mar 28. doi:10.1186/s12885-017-3226-4.

Menck K, Wlochowicz D, Wachter A, et al. High-Throughput Profiling of Colorectal Cancer Liver Metastases Reveals Intra- and Inter-Patient Heterogeneity in the EGFR and WNT Pathways Associated with Clinical Outcome. *Cancers (Basel)*. 2022, 14(9):2084. Published 2022 Apr 21. doi:10.3390/cancers14092084.

Messersmith WA. NCCN Guidelines Updates: Management of Metastatic Colorectal Cancer. *J Natl Compr Canc Netw*. 2019, 17(5.5):599-601. doi:10.6004/jnccn.2019.5014.

Ministério da Saúde. Instituto Nacional de Câncer José de Alencar Gomes da Silva. Estatísticas de câncer. 2020. Disponível em: <URL:<https://www.inca.gov.br/numeros-de-cancer>>. [2021 dez 15].

Mitsa G, Guo Q, Goncalves C, et al. A Non-Hazardous Deparaffinization Protocol Enables Quantitative Proteomics of Core Needle Biopsy-Sized Formalin-Fixed and Paraffin-Embedded (FFPE) Tissue Specimens. *Int J Mol Sci*. 2022, 23(8):4443. Published 2022 Apr 18. doi:10.3390/ijms23084443.

Mizuno R, Kawada K, Itatani Y, Ogawa R, Kiyasu Y, Sakai Y. The Role of Tumor-Associated Neutrophils in Colorectal Cancer. *Int J Mol Sci*. 2019, 20(3):529. Published 2019 Jan 27. doi:10.3390/ijms20030529.

Montemagno C, Pagès G. Resistance to Anti-angiogenic Therapies: A Mechanism Depending on the Time of Exposure to the Drugs. *Front Cell Dev Biol*. 2020, 8:584. Published 2020 Jul 7. doi:10.3389/fcell.2020.00584.

Morgan RG, Mortenson E, Williams AC. Targeting LGR5 in Colorectal Cancer: therapeutic gold or too plastic?. *Br J Cancer*. 2018, 118(11):1410-1418. doi:10.1038/s41416-018-0118-6.

- Mukund K, Syulyukina N, Ramamoorthy S, Subramaniam S. Right and left-sided colon cancers - specificity of molecular mechanisms in tumorigenesis and progression. *BMC Cancer*. 2020, 20(1):317. Published 2020 Apr 15. doi:10.1186/s12885-020-06784-7.
- Naba A, Clauser KR, Whittaker CA, Carr SA, Tanabe KK, Hynes RO. Extracellular matrix signatures of human primary metastatic colon cancers and their metastases to liver. *BMC Cancer*. 2014, 14:518. Published 2014 Jul 18. doi:10.1186/1471-2407-14-518.
- Nakanishi Y, Diaz-Meco MT, Moscat J. Serrated Colorectal Cancer: The Road Less Travelled?. *Trends Cancer*. 2019, 5(11):742-754. doi:10.1016/j.trecan.2019.09.004.
- Nguyen V, Griss J. scAnnotatR: framework to accurately classify cell types in single-cell RNA-sequencing data. *BMC Bioinformatics*. 2022;23(1):44. Published 2022 Jan 17. doi:10.1186/s12859-022-04574-5.
- Niida A, Mimori K, Shibata T, Miyano S. Modeling colorectal cancer evolution. *J Hum Genet*. 2021, 66(9):869-878. doi:10.1038/s10038-021-00930-0.
- Nowell PC. The clonal evolution of tumor cell populations. *Science*. 1976, 194(4260):23-28. doi:10.1126/science.959840.
- O'Neill JR. An Overview of Mass Spectrometry-Based Methods for Functional Proteomics. *Methods Mol Biol*. 2019, 1871:179-196. doi:10.1007/978-1-4939-8814-3_13.
- Ormseth B, Onuma A, Zhang H, Tsung A. The Hepatic Pre-Metastatic Niche. *Cancers (Basel)*. 2022, 14(15):3731. Published 2022 Jul 31. doi:10.3390/cancers14153731.
- Ouyang R, Li Z, Peng P, Zhang J, Liu J et al. Exploration of the relationship between tumor mutation burden and immune infiltrates in colon adenocarcinoma. *Int J Med Sci*. 2021 Jan 1, 18(3):685-694.
- Ozawa S, Kato Y, Komori R et al. BRAK/CXCL14 expression suppresses tumor growth in vivo in human oral carcinoma cells. *Biochem Biophys Res Commun*. 2006 Sep 22;348(2):406-12. doi: 10.1016/j.bbrc.2006.07.070
- Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021, 372:n71. Published 2021 Mar 29. doi:10.1136/bmj.n71.
- Patel JN, Fong MK, Jagosky M. Colorectal Cancer Biomarkers in the Era of Personalized Medicine. *J Pers Med*. 2019, 9(1):3. Published 2019 Jan 14. doi:10.3390/jpm9010003.
- Paton EL, Turner JA, Schlaepfer IR. Overcoming Resistance to Therapies Targeting the MAPK Pathway in BRAF-Mutated Tumours. *J Oncol*. 2020, 2020:1079827. Published 2020 Jan 3. doi:10.1155/2020/1079827.
- Pedrosa L, Esposito F, Thomson TM, Maurel J. The Tumor Microenvironment in Colorectal Cancer Therapy. *Cancers (Basel)*. 2019, 11(8):1172. Published 2019 Aug 14. doi:10.3390/cancers11081172.

Penney KL, Banbury BL, Bien S, et al. Genetic Variant Associated With Survival of Patients With Stage II-III Colon Cancer. *Clin Gastroenterol Hepatol*. 2020, 18(12):2717-2723.e3. doi:10.1016/j.cgh.2019.11.046.

Piawah S, Venook AP. Targeted therapy for colorectal cancer metastases: A review of current methods of molecularly targeted therapy and the use of tumor biomarkers in the treatment of metastatic colorectal cancer. *Cancer*. 2019, 125(23):4139-4147. doi:10.1002/cncr.32163.

Poturnajova M, Furielova T, Balintova S, Schmidtova S, Kucerovala L, Matuskova M. Molecular features and gene expression signature of metastatic colorectal cancer (Review). *Oncol Rep*. 2021, 45(4):10. doi:10.3892/or.2021.7961.

Prahallad A, Sun C, Huang S, et al. Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature*. 2012, 483(7387):100-103. Published 2012 Jan 26. doi:10.1038/nature10868.

Pranteda A, Piastra V, Stramucci L, Fratantonio D, Bossi G. The p38 MAPK Signaling Activation in Colorectal Cancer upon Therapeutic Treatments. *Int J Mol Sci*. 2020, 21(8):2773.

Price TJ, Peeters M, Kim TW, et al. Panitumumab versus cetuximab in patients with chemotherapy-refractory wild-type KRAS exon 2 metastatic colorectal cancer (ASPECCT): a randomised, multicentre, open-label, non-inferiority phase 3 study. *Lancet Oncol*. 2014, 15(6):569-579. doi:10.1016/S1470-2045(14)70118-4.

Punt CJ, Koopman M, Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat Rev Clin Oncol*. 2017, 14(4):235-246. doi:10.1038/nrclinonc.2016.171.

Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015, 43(7):e47. doi:10.1093/nar/gkv007.

Rodriguez H, Zenklusen JC, Staudt LM, Doroshow JH, Lowy DR. The next horizon in precision oncology: Proteogenomics to inform cancer diagnosis and treatment. *Cell*. 2021, 184(7):1661-1670. doi:10.1016/j.cell.2021.02.055.

Rozanova S, Barkovits K, Nikolov M, Schmidt C, Urlaub H, Marcus K. Quantitative Mass Spectrometry-Based Proteomics: An Overview. *Methods Mol Biol*. 2021, 2228:85-116. doi:10.1007/978-1-0716-1024-4_8.

Sagaert X, Vanstapel A, Verbeek S. Tumor Heterogeneity in Colorectal Cancer: What Do We Know So Far?. *Pathobiology*. 2018, 85(1-2):72-84.

Sainson RCA, Thotakura AK, Kosmac M, et al. An Antibody Targeting ICOS Increases Intratumoral Cytotoxic to Regulatory T-cell Ratio and Induces Tumor Regression. *Cancer Immunol Res*. 2020, 8(12):1568-1582. doi:10.1158/2326-6066.CIR-20-0034.

Samaržija I. Wnt Signaling Pathway Is among the Drivers of Liver Metastasis. *Livers*. 2021, 1(4):180-200. doi:10.3390/livers1040015.

Sarvestani SK, DeHaan RK, Miller PG, et al. A Tissue Engineering Approach to Metastatic Colon Cancer. *iScience*. 2020, 23(11):101719. Published 2020 Oct 20. doi:10.1016/j.isci.2020.101719.

Sayagués JM, Corchete LA, Gutiérrez ML, et al. Genomic characterization of liver metastases from colorectal cancer patients. *Oncotarget*. 2016, 7(45):72908-72922. doi:10.18632/oncotarget.12140.

Schubert M, Klinger B, Klünemann M, et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat Commun*. 2018;9(1):20. Published 2018 Jan 2. doi:10.1038/s41467-017-02391-6.

Schwartz LH, Litière S, de Vries E, et al. RECIST 1.1-Update and clarification: From the RECIST committee. *Eur J Cancer*. 2016, 62:132-137. doi:10.1016/j.ejca.2016.03.081.

Shangkuan WC, Lin HC, Chang YT et al. Risk analysis of colorectal cancer incidence by gene expression analysis. *PeerJ*. 2017 Feb 15;5:e3003. doi: 10.7717/peerj.3003.

Shi R, Chen W, Yang B, et al. Prediction of KRAS, NRAS and BRAF status in colorectal cancer patients with liver metastasis using a deep artificial neural network based on radiomics and semantic features. *Am J Cancer Res*. 2020, 10(12):4513-4526. Published 2020 Dec 1.

Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin*. 2020, 70(3):145-164. doi:10.3322/caac.21601.

Simon K. Colorectal cancer development and advances in screening. *Clin Interv Aging*. 2016, 11:967-976. Published 2016 Jul 19. doi:10.2147/CIA.S109285.

Stanisavljević L, Myklebust MP, Leh S, Dahl O. LGR5 and CD133 as prognostic and predictive markers for fluoropyrimidine-based adjuvant chemotherapy in colorectal cancer. *Acta Oncol*. 2016, 55(12):1425-1433. doi:10.1080/0284186X.2016.1201215.

Sturm G, Finotello F, List M. Immunedeconv: An R Package for Unified Access to Computational Methods for Estimating Immune Cell Fractions from Bulk RNA-Sequencing Data. *Methods Mol Biol*. 2020;2120:223-232. doi:10.1007/978-1-0716-0327-7_16.

Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021, 71(3):209-249. doi:10.3322/caac.21660.

Tabernero J, Grothey A, Van Cutsem E, et al. Encorafenib Plus Cetuximab as a New Standard of Care for Previously Treated BRAF V600E-Mutant Metastatic Colorectal Cancer: Updated Survival Results and Subgroup Analyses from the BEACON Study. *J Clin Oncol*. 2021, 39(4):273-284. doi:10.1200/JCO.20.02088.

Taieb J, André T, Auclin E. Refining adjuvant therapy for non-metastatic colon cancer, new standards and perspectives. *Cancer Treat Rev*. 2019, 75:1-11. doi:10.1016/j.ctrv.2019.02.002.

Teng S, Li YE, Yang M, et al. Tissue-specific transcription reprogramming promotes liver metastasis of colorectal cancer. *Cell Res*. 2020, 30(1):34-49. doi:10.1038/s41422-019-0259-z.

Tessema M, Klinge DM, Yingling CM et al. Re-expression of CXCL14, a common target for epigenetic silencing in lung cancer, induces tumor necrosis. *Oncogene*. 2010 Sep 16;29(37):5159-70. doi: 10.1038/onc.2010.255

Thanikachalam K, Khan G. Colorectal Cancer and Nutrition. *Nutrients*. 2019, 11(1):164. Published 2019 Jan 14. doi:10.3390/nu11010164.

Tiwari A, Saraf S, Verma A, Panda PK, Jain SK. Novel targeting approaches and signaling pathways of colorectal cancer: An insight. *World J Gastroenterol*. 2018, 24(39):4428-4435. doi:10.3748/wjg.v24.i39.4428.

Tsilimigras DI, Brodt P, Clavien PA, et al. Liver metastases. *Nat Rev Dis Primers*. 2021, 7(1):27. Published 2021 Apr 15. doi:10.1038/s41572-021-00261-6.

Tsilimigras DI, Ntanasis-Stathopoulos I, Bagante F, et al. Clinical significance and prognostic relevance of KRAS, BRAF, PI3K and TP53 genetic mutation analysis for resectable and unresectable colorectal liver metastases: A systematic review of the current evidence. *Surg Oncol*. 2018, 27(2):280-288. doi:10.1016/j.suronc.2018.05.012.

Vakiani E, Janakiraman M, Shen R, et al. Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J Clin Oncol*. 2012, 30(24):2956-2962. doi:10.1200/JCO.2011.38.2994.

Valenzuela G, Canepa J, Simonetti C, Solo de Zaldívar L, Marcelain K, González-Montero J. Consensus molecular subtypes of colorectal cancer in clinical practice: A translational approach. *World J Clin Oncol*. 2021, 12(11):1000-1008. doi:10.5306/wjco.v12.i11.1000.

van Huizen NA, Coebergh van den Braak RRJ, Doukas M, Dekker LJM, IJzermans JNM, Luidert TM. Up-regulation of collagen proteins in colorectal liver metastasis compared with normal liver tissue. *J Biol Chem*. 2019, 294(1):281-289. doi:10.1074/jbc.RA118.005087.

Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, et al. Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. *Cell*. 2019, 177(4):1035-1049.e19.

Vasilogianni AM, Al-Majdoub ZM, Achour B, Peters SA, Rostami-Hodjegan A, Barber J. Proteomics of colorectal cancer liver metastasis: A quantitative focus on drug elimination and pharmacodynamics effects. *Br J Clin Pharmacol*. 2022, 88(4):1811-1823. doi:10.1111/bcp.15098.

Vera R, González-Flores E, Rubio C, et al. Multidisciplinary management of liver metastases in patients with colorectal cancer: a consensus of SEOM, AEC, SEOR, SERVEI, and SEMNIM. *Clin Transl Oncol*. 2020, 22(5):647-662. doi:10.1007/s12094-019-02182-z.

Villalba M, Evans SR, Vidal-Vanaclocha F, Calvo A. Role of TGF- β in metastatic colon cancer: it is finally time for targeted therapy. *Cell Tissue Res*. 2017, 370(1):29-39. doi:10.1007/s00441-017-2633-9.

Voß H, Wurlitzer M, Smit DJ, et al. Differential regulation of extracellular matrix proteins in three recurrent liver metastases of a single patient with colorectal cancer. *Clin Exp Metastasis*. 2020, 37(6):649-656. doi:10.1007/s10585-020-10058-8.

Wang L, Jin J, Zhou Y, et al. EMX2 is epigenetically silenced and suppresses epithelial-mesenchymal transition in human esophageal adenocarcinoma. *Oncol Rep*. 2019, 42(5):2169-2178. doi:10.3892/or.2019.7284.

Wang X, Zhang H, Chen X. Drug resistance and combating drug resistance in cancer. *Cancer Drug Resist*. 2019, 2(2):141-160. doi:10.20517/cdr.2019.10.

Wang W, Huang P, Zhang L et al. Antitumor efficacy of C-X-C motif chemokine ligand 14 in hepatocellular carcinoma in vitro and in vivo. *Cancer Sci*. 2013 Nov;104(11):1523-31. doi: 10.1111/cas.12279

Weiser MR. AJCC 8th Edition: Colorectal Cancer. *Ann Surg Oncol*. 2018, 25(6):1454-1455. doi: 10.1245/s10434-018-6462-1.

Wilson SM, Adson MA. Surgical treatment of hepatic metastases from colorectal cancers. *Arch Surg* 1976; 111:330-4.

Wong GYM, Diakos C, Hugh TJ, Molloy MP. Proteomic Profiling and Biomarker Discovery in Colorectal Liver Metastases. *Int J Mol Sci*. 2022, 23(11):6091. Published 2022 May 29. doi:10.3390/ijms23116091.

Wu CC, Wang JH, Lin PC, et al. Tumor sidedness and efficacy of first-line therapy in patients with RAS/BRAF wild-type metastatic colorectal cancer: A network meta-analysis. *Crit Rev Oncol Hematol*. 2020, 145:102823. doi:10.1016/j.critrevonc. 2019.102823.

Wu M, Ma M, Tan Z, Zheng H, Liu X. Neutrophil: A New Player in Metastatic Cancers. *Front Immunol*. 2020, 11:565165. Published 2020 Sep 24. doi:10.3389/fimmu.2020.565165.

Wu W, Wang Q, Yin F, et al. Identification of proteomic and metabolic signatures associated with chemoresistance of human epithelial ovarian cancer. *Int J Oncol*. 2016, 49(4):1651-1665. doi:10.3892/ijo.2016.3652.

Xie YH, Chen YX, Fang JY. Comprehensive review of targeted therapy for colorectal cancer. *Signal Transduct Target Ther*. 2020, 5(1):22. Published 2020 Mar 20. doi:10.1038/s41392-020-0116-z.

Yang D, Wang LL, Dong TT, et al. Progranulin promotes colorectal cancer proliferation and angiogenesis through TNFR2/Akt and ERK signaling pathways. *Am J Cancer Res*. 2015, 5(10):3085-3097. Published 2015 Sep 15.

Yeh CC, Shih LJ, Chang JL, et al. Synchronous vascular endothelial growth factor protein profiles in both tissue and serum identify metastasis and poor survival in colorectal cancer. *Sci Rep*. 2019, 9(1):4228. Published 2019 Mar 12. doi:10.1038/s41598-019-40862-6.

Yeoh Y, Low TY, Abu N, Lee PY. Regulation of signal transduction pathways in colorectal cancer: implications for therapeutic resistance. *PeerJ*. 2021, 9:e12338. Published 2021 Oct 22. doi:10.7717/peerj.12338.

Yuan Z, Liang X, Zhan Y, et al. Targeting CD133 reverses drug-resistance via the AKT/NF- κ B/MDR1 pathway in colorectal cancer. *Br J Cancer*. 2020, 122(9):1342-1353. doi:10.1038/s41416-020-0783-0.

Yuan Z, Liang X, Zhan Y, et al. Targeting CD133 reverses drug-resistance via the AKT/NF- κ B/MDR1 pathway in colorectal cancer. *Br J Cancer*. 2020, 122(9):1342-1353. doi:10.1038/s41416-020-0783-0.

Yue D, Li H, Che J, et al. EMX2 Is a Predictive Marker for Adjuvant Chemotherapy in Lung Squamous Cell Carcinomas. *PLoS One*. 2015, 10(7):e0132134. Published 2015 Jul 1. doi:10.1371/journal.pone.0132134.

Zarour LR, Anand S, Billingsley KG, et al. Colorectal Cancer Liver Metastasis: Evolving Paradigms and Future Directions. *Cell Mol Gastroenterol Hepatol*. 2017, 3(2):163-173. Published 2017 Jan 20. doi:10.1016/j.jcmgh.2017.01.006.

Zhang B, Wang J, Wang X, et al. Proteogenomic characterization of human colon and rectal cancer. *Nature*. 2014, 513(7518):382-387. doi:10.1038/nature13438.

Zhang C, Mei Z, Pei J, Abe M, Zeng X, Huang Q, et al. A Modified Tumor-Node-Metastasis Classification for Primary Operable Colorectal Cancer. *JNCI Cancer Spectr*. 2020, 5(1):pkaa093.

Zhang T, Guo J, Gu J, et al. Identifying the key genes and microRNAs in colorectal cancer liver metastasis by bioinformatics analysis and in vitro experiments. *Oncol Rep*. 2019, 41(1):279-291. doi:10.3892/or.2018.6840.

Zhang T, Yuan K, Wang Y, et al. Identification of Candidate Biomarkers and Prognostic Analysis in Colorectal Cancer Liver Metastases. *Front Oncol*. 2021, 11:652354. Published 2021 Aug 4. doi:10.3389/fonc.2021.652354.

Zhang G, Luo X, Zhang W. CXCL-13 Regulates Resistance to 5-Fluorouracil in Colorectal Cancer. *Cancer Res Treat*. 2020 Apr; 52(2): 622–633. doi: 10.4143/crt.2019.593

Zhao P, Zhang Z. TNF- α promotes colon cancer cell migration and invasion by upregulating TROP-2. *Oncol Lett*. 2018, 15(3):3820-3827. doi:10.3892/ol.2018.7735.

Zheng Z, Yu T, Zhao X, Gao X, Zhao Y, Liu G. Intratumor heterogeneity: A new perspective on colorectal cancer research. *Cancer Med*. 2020, 9(20):7637-7645.

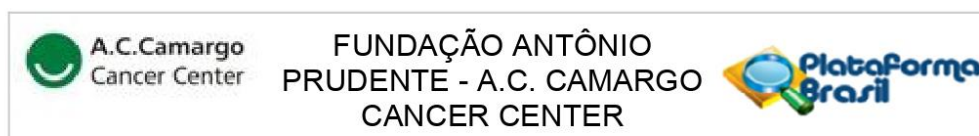
Zhou H, Liu Z, Wang Y, et al. Colorectal liver metastasis: molecular mechanism and interventional therapy. *Signal Transduct Target Ther*. 2022, 7(1):70. Published 2022 Mar 4. doi:10.1038/s41392-022-00922-2.

Zhou SN, Pan WT, Pan MX, et al. Comparison of Immune Microenvironment Between Colon and Liver Metastatic Tissue in Colon Cancer Patients with Liver Metastasis. *Dig Dis Sci*. 2021, 66(2):474-482. doi:10.1007/s10620-020-06203-8.

Zhu Y, Yu J, Wang S, Lu R, Wu J, Jiang B. Overexpression of CD133 enhances chemoresistance to 5-fluorouracil by activating the PI3K/Akt/p70S6K pathway in gastric cancer cells. *Oncol Rep*. 2014, 32(6):2437-2444. doi:10.3892/or.2014.3488.

8 ANEXOS

Anexo 1 – Carta de aprovação do Comitê de Ética em Pesquisa (CEP)



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Identificação de moduladores da sinalização celular no câncer de cólon e metástases hepáticas como ferramenta para descoberta de alvos terapêuticos

Pesquisador: Fabio Albuquerque Marchi

Área Temática: Genética Humana:
(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP;);

Versão: 1

CAAE: 59038622.0.0000.5432

Instituição Proponente: FUNDACAO ANTONIO PRUDENTE

Patrocinador Principal: FUNDACAO DE AMPARO A PESQUISA DO ESTADO DE SAO PAULO

DADOS DO PARECER

Número do Parecer: 5.480.477

Apresentação do Projeto:

As informações elencadas nos itens "Apresentação do Projeto", "Objetivo da Pesquisa", "Avaliação dos Riscos e Benefícios" foram retiradas do documento "PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1948741.pdf", gerado pela Plataforma Brasil em 26/05/2022; e do documento "Projeto_Daniela_Mestrado.pdf", postado na PB em 20/05/2022.

Resumo:

O câncer de cólon é um dos tipos de câncer com as maiores taxas mundiais de incidência e mortalidade. O elevado número de óbitos está relacionado, principalmente, ao fato de que aproximadamente metade dos pacientes acabam desenvolvendo metástases hepáticas durante o curso da doença, o que decai drasticamente a expectativa de vida desses indivíduos. Apesar da ressecção cirúrgica das metástases ser realizada com intenção curativa, com uma taxa de sobrevida relativa de 5 anos ao redor de 40%, grande parte dos pacientes não são elegíveis ao tratamento, recorrendo às abordagens paliativas com limitadas opções terapêuticas devido ao perfil de resistência que as células tumorais adquirem ao longo do tratamento ofertado. Embora o rastreamento molecular das metástases tenha possibilitado a estratificação de pacientes de acordo com a eficiência das terapias disponíveis, ainda é limitado o conhecimento sobre o perfil global de

Endereço: Rua Professor Antonio Prudente, 211 º, Liberdade º, Atrium - Subsolo
Bairro: Liberdade **CEP:** 01.509-900
UF: SP **Município:** SAO PAULO
Telefone: (11)2189-5020 **Fax:** (11)2189-5020 **E-mail:** cep_accamargo@accamargo.org.br



**FUNDAÇÃO ANTÔNIO
PRUDENTE - A.C. CAMARGO
CANCER CENTER**



Continuação do Parecer: 5.480.477

Infraestrutura	Infraestrutura_Assinado.pdf	26/05/2022 15:12:09	Fabio Albuquerque Marchi	Aceito
Outros	Declaracao_Biobanco.pdf	26/05/2022 15:11:44	Fabio Albuquerque Marchi	Aceito
Outros	Declaracao_sobre_os_dados_coletados.pdf	20/05/2022 11:03:11	Fabio Albuquerque Marchi	Aceito
Outros	Declaracao_Sobre_a_Publicacao_dos_Dados_da_Pesquisa.pdf	20/05/2022 10:54:05	Fabio Albuquerque Marchi	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	Termo_de_Dispensa_do_Termo_de_Consentimento_Livre_e_Esclarecido.pdf	20/05/2022 10:53:12	Fabio Albuquerque Marchi	Aceito
Orçamento	Orcamento_Financeiro_Detalhado.pdf	20/05/2022 10:52:53	Fabio Albuquerque Marchi	Aceito
Outros	Identificacao_dos_CVs_na_Plataforma_Lattes.pdf	20/05/2022 10:52:40	Fabio Albuquerque Marchi	Aceito
Outros	Formulario_de_submissao_ao_CEP.pdf	20/05/2022 10:52:23	Fabio Albuquerque Marchi	Aceito
Declaração do Patrocinador	Termo_de_Compromisso_do_Pesquisador.pdf	20/05/2022 10:51:58	Fabio Albuquerque Marchi	Aceito
Outros	Declaracao_Patologia.pdf	20/05/2022 10:51:36	Fabio Albuquerque Marchi	Aceito
Declaração de concordância	Declaracao_OncoClinica.pdf	20/05/2022 10:51:03	Fabio Albuquerque Marchi	Aceito
Declaração de Manuseio Material Biológico / Biorepositório / Biobanco	Declaracao_Sobre_o_Uso_e_Destino_do_Material_Biologico.pdf	20/05/2022 10:50:09	Fabio Albuquerque Marchi	Aceito
Cronograma	Cronograma_de_estudo.pdf	20/05/2022 10:49:52	Fabio Albuquerque Marchi	Aceito
Projeto Detalhado / Brochura Investigador	Projeto_Daniela_Mestrado.pdf	20/05/2022 10:42:32	Fabio Albuquerque Marchi	Aceito
Folha de Rosto	folhaDeRosto_PB_Assinada.pdf	20/05/2022 10:41:38	Fabio Albuquerque Marchi	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Endereço: Rua Professor Antonio Prudente, 211 º, Liberdade º Atrium - Subsolo
Bairro: Liberdade **CEP:** 01.509-900
UF: SP **Município:** SAO PAULO
Telefone: (11)2189-5020 **Fax:** (11)2189-5020 **E-mail:** cep_accamargo@accamargo.org.br