

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE BIOCÊNCIAS

DEPARTAMENTO DE GENÉTICA

**ANÁLISE DO POSSÍVEL PAPEL DE TUBBY-LIKE PROTEIN 3 (TULP3)
COMO BIOMARCADOR EM TUMORES DO TRATO GASTRINTESTINAL**

Ivaine Taís Sauthier Sartor

Tese submetida ao Programa de Pós-
Graduação em Genética e Biologia
Molecular da UFRGS como requisito
parcial para a obtenção do grau de
Doutor

Orientadora: Profa. Dra. Patricia Ashton-Prolla

Co-orientadora: Profa. Dra. Mariana Recamonde-Mendoza

Porto Alegre, Fevereiro de 2019.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE BIOCÊNCIAS

DEPARTAMENTO DE GENÉTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

Este trabalho foi desenvolvido no Laboratório de Medicina Genômica do Centro de Pesquisa Experimental do Hospital de Clínicas de Porto Alegre e no Núcleo de Pesquisas em Oncologia da Universidade Federal do Pará e foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Fundo de Incentivo à Pesquisa e Eventos (FIPE-HCPA).

Agradecimentos

Agradeço, em especial à Professora Patricia Ashton-Prolla, pela oportunidade de realizar este trabalho, por todos os ensinamentos, pela ousadia em aceitar novos desafios, e pelo exemplo de dedicação à pesquisa.

Agradeço também à Professora Mariana Recamonde-Mendoza pelo fundamental auxílio nas análises de bioinformática e pela agradável companhia.

Ao Professor Alessandro Osvaldt, à Professora Raquel Rivero e à Simone Machado e pela oportunidade de aprender tanto com todos vocês.

Agradeço também à Bárbara Alemar pela convivência e aprendizado e à Cleandra Gregório, pela amizade e por todo conhecimento que adquirimos. Os dias ao lado de vocês foram de muito trabalho, mas repletos de humor.

Agradeço também aos Professores do NPO-UFGPA Samia Demakchi, Paulo Assumpção, Andrea e Sidney dos Santos que me acolheram na minha estada em Belém. Em especial, agradeço à Ana Anaissi que, além de contribuir para esse estudo, tornou-se uma grande amiga.

Aos colegas do Laboratório de Medicina Genômica por toda ajuda e convivência, em especial ao Gabriel Macedo, à Clévia Rosset e à Professora Fernanda Vianna.

Aos Professores do Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS pela excelência no ensino.

Ao Elmo Cardoso e Everaldo Almeida, pela dedicação e por todo auxílio prestado no PPGBM e no HCPA. Muito obrigada!

Aos órgãos de fomento à pesquisa, em especial ao CNPq pela bolsa de estudos concedida e ao Fundo de Incentivo à Pesquisa do Hospital de Clínicas de Porto Alegre, que viabilizou a realização deste trabalho.

A todos os pacientes que aceitaram participar de estudos de científicos.

Aos meus queridos colegas do Projeto Mais TMO do Hospital Moinhos de Vento pela agradável convivência.

Agradeço às minhas amigas de toda hora, Rafaela Haefliger, Patrícia Silva, Camilla Bittar, Liliane Todeschini, Gabriela Zuqueto e Maria José Dornelles que tornaram meus dias leves e divertidíssimos ao lado de vocês.

Por fim, agradeço aos meus pais, Ivo e Elaine, meus grandes exemplos de caráter e coragem, por todo amor e apoio ao longo da minha vida. Aos meus companheiros de quatro patas, Japi e Nica, pelo carinho. À minha querida tia Geni Haefliger e sua família pela companhia pra lá de especial. Sou muito mais feliz com vocês ao meu lado.

Sumário

| | |
|---|-----|
| Resumo | 10 |
| Abstract | 12 |
| Capítulo I | 14 |
| 1. Câncer e biomarcadores | 15 |
| 1.1 Câncer gástrico | 16 |
| 1.2 Câncer colorretal | 20 |
| 1.3 Câncer pancreático | 24 |
| 1.4 Biomarcadores tumorais | 27 |
| 2. Biologia molecular e análise computacional | 35 |
| 3. Tubby-Like Protein 3 (TULP3) | 38 |
| 4. Justificativa | 41 |
| 5. Objetivos | 42 |
| 5.1 Objetivo geral | 42 |
| 5.2 Objetivos específicos | 42 |
| Capítulo II | 43 |
| Capítulo III | 77 |
| Capítulo IV | 104 |
| Capítulo V | 134 |
| Capítulo VI | 147 |
| Capítulo VII | 154 |
| Capítulo VIII | 156 |
| Anexos | 158 |
| Referências | 162 |

Lista de Abreviaturas

| | |
|--|---|
| 5-FU: 5-fluorouracil | FAP: Polipose adenomatosa familiar |
| ACRG: <i>Asian Cancer Research Group</i> | FISH: <i>Fluorescent in situ Hybridization</i> |
| ADP: Adenocarcinoma ductal pancreático | GC: Genomicamente estável |
| AJCC: <i>American Joint Committee on Cancer</i> | GEO: <i>Gene Expression Omnibus</i> |
| APC: <i>Adenomatosis Polyposis Coli</i> | GI: Gastrointestinal |
| ARACNe: <i>Algorithm for Reconstruction of Accurate Cellular Network</i> | GIM: Metaplasia intestinal completa |
| BRAF: <i>B-Raf proto-oncogene</i> | GIM: Metaplasia intestinal gástrica |
| BRCA2: <i>Breast Cancer type 2</i> | GPCR: Receptores acoplados à proteína G |
| CA19-9: Antígeno carboidrato 19-9 | GRCh38: <i>Genome Reference Consortium Human Build 38</i> |
| CCR: Câncer colorretal | HER2: <i>Erb-B2 Receptor Tyrosine Kinase 2 (ERBB2)</i> |
| CDH1: <i>Cadherin 1</i> | HH: Via de sinalização Hedgehog |
| CDKN2: <i>Cyclin Dependent Kinase Inhibitor 2A</i> | HNPCC: Câncer colorretal hereditário não-polipoide |
| CDX1: <i>Caudal Type Homeobox 1</i> | ICG: Câncer gástrico do tipo intestinal |
| CDX2: <i>Caudal Type Homeobox 2</i> | IFT-A: Transporte ciliar intraflagelar do complexo A |
| CEA: Antígeno carcinoembrionário | IHQ: Imunohistoquímica |
| CG: Câncer gástrico | IIM: Metaplasia intestinal incompleta |
| ChG: Gastrite crônica | IPMN: Neoplasia intraductal papilífera mucinosa |
| CIN: Instabilidade cromossômica | KRAS: <i>Kirsten Rat Sarcoma Viral Oncogene Homolog</i> |
| COAD: Adenocarcinoma de cólon | MAPK: |
| DCG: Câncer gástrico do tipo difuso | |
| EBV: vírus Epstein-Barr | |
| FAMMM: Síndrome familiar atípica de múltiplos nevos melanocíticos | |

MCN: Neoplasia cística mucinosa

MLH1: *MutL Homolog 1*

MMR: *Mismatch Repair system*

MRA: *Master Regulator Analysis*

MSH2: *MutS Homolog 2*

MSH6: *MutS Homolog 6*

MSI: instabilidade microssatélite

MSS/EMT: estabilidade microssatélite com transição epitélio-mesenquimal

MSS/TP53-: estabilidade microssatélite com proteína tumoral 53 inativa

MSS/TP53+: estabilidade microssatélite com proteína tumoral 53 ativa

NCI: *National Cancer Institute*

NHGRI: *National Human Genome Research Institute*

NRAS: *Neuroblastoma RAS Viral Oncogene Homolog*

NT: tecido não-tumoral adjacente

OMS: Organização Mundial da Saúde

PanIN: neoplasia intraepitelial pancreática

PI3K: fosfatidilinositol-3 quinase

PIP₂: fosfoinosítideo de membrana

PMS2: *PMS1 Homolog 2, Mismatch Repair Protein*

PTEN: *Phosphatase And Tensin Homolog*

READ: adenocarcinoma de reto

RNA-seq: sequenciamento de DNA complementar seguido de quantificação

SMAD/DPC4:

SNPs: polimorfismos de base única

TCGA: *The Cancer Genome Atlas*

ToGA: *Trastuzumabe for Gastric Cancer*

TP53: *Tumor Protein P53*

TULP3: *Tubby-like Protein 3*

Lista de Figuras

- Figura 1.** Estimativas mundiais de novos casos para o ano de 2018. Os sítios tumorais que apresentam maiores taxas são demonstrados no gráfico (Globocan, acesso em 28 de Janeiro de 2020)..... 15
- Figura 2.** Estimativas mundiais de mortalidade para o ano de 2018. Os sítios tumorais que apresentam maiores taxas de mortalidade são demonstrados no gráfico (Globocan, acesso em 28 de Janeiro de 2020)..... 16
- Figura 3.** Camadas histológicas que compõe o estômago (American Cancer Society, acesso em 31 de Janeiro de 2019)..... 17
- Figura 4.** Classificação histológica do adenocarcinoma gástrico, segundo Laurén. (A) tipo intestinal, (B) tipo difuso e (C) tipo misto. (Chong *et al.*, 2014)..... 18
- Figura 5.** Classificação dos subtipos moleculares consenso de tumores colorretais (CMS). (a) Rede dos subtipos CMS1-4, relacionada aos seis subtipos (A-F) previamente identificados pelo Consórcio para subtipagem do CCR (CRCSC). (b) Distribuição por amostra de cada uma das classificações (A-F) agrupados pelos clusters dos subtipos CMS. As cores dentro de cada linha representam subtipos diferentes. (c) Rede de pacientes diagnosticados com CCR. (d) Distribuição final dos grupos CMS1–4 (cores sólidas), amostras 'mistas' (cores em gradiente), e amostras indeterminadas (cor cinza). 23
- Figura 6.** Modelo de progressão do câncer pancreático. Alterações genéticas são acumuladas durante a transformação do ducto normal em lesões intraepiteliais ao câncer invasivo (Hruban *et al.*, 2001)..... 27
- Figura 7.** Domínios de ligação de TULP3. (a) TULP3 se associa ao core IFT-A através da região N-terminal. TULP3-(IFT-A) aprisionam GPCRs ciliares à vesícula através da ligação de TULP3 com PIP₂ são transportados via IFT-A de forma retrógrada e anterógrada no cílio primário. (b) Domínio tubby de TULP3 é ancorado à PIP₂, na membrana. A ativação da GPCR causa a hidrólise de PIP₂ e a liberação de TULP3 no

citoplasma. (IFT-A) transporte intraflagelar do complexo A. (IP₃) inositol triphosphate.
(DAG) diacylglicerol. (PCL) phospholipase C. (Mukhopadhyay & Jackson, 2011,
modificado)..... 39

Resumo

Os tumores gastrointestinais são reconhecidos mundialmente como doenças heterogêneas e agressivas que apresentam um prognóstico desfavorável tanto para os pacientes que não realizaram a cirurgia de ressecção quanto àqueles que passaram pelo procedimento de ressecção do órgão. A expressão de RNAm e/ou proteínas tumorais específicas, ou ainda, alguma alteração genética, identificada em diferentes neoplasias, aliadas às informações clínicas, podem conferir importantes fontes de informações para o correto diagnóstico, prognóstico ou predição a resposta de tratamentos. Em especial, o RNAm apresenta papel fundamental na progressão e manutenção tumoral e, além disso, tem maior potencial de refletir os fenótipos celulares quando comparados a alterações no DNA. A hipótese desse estudo baseou-se em achados prévios, nos quais o fator de transcrição *TULP3* foi apontado como regular mestre da carcinogênese pancreática. Neste estudo verificou-se que os elevados níveis de expressão de *TULP3* apresentaram associação significativa com um desfecho clínico desfavorável. Ainda, esses valores aumentados de expressão gênica não possuíam qualquer valor prognóstico em outros tipos tumorais, como cânceres de mama, ovário e pulmão, indicando a especificidade como marcador prognóstico no adenocarcinoma ductal pancreático. O pâncreas, assim como outros órgãos do trato gastrintestinal, em especial o estômago e intestinos, possuem a mesma origem embrionária, o endoderma. Sendo esses órgãos de mesma origem, nossa hipótese é de que o fator de transcrição *TULP3* possa apresentar uma função como biomarcador em neoplasias do sistema gastrintestinal. Este estudo buscou investigar o perfil de expressão *TULP3* em doenças gastrintestinais, utilizando dados disponíveis em bancos de dados públicos e amostras de tecidos de tumores do trato gastrintestinal, em especial, os tumores pancreáticos, gástricos e colorretais. Este trabalho associou, pela primeira vez, a expressão de *TULP3* à carcinogênese gástrica. Através de análises bioinformáticas utilizando dados de

expressão gênica disponíveis em bancos de dados públicos os níveis elevados de *TULP3* foram associados a um prognóstico desfavorável para o tipo difuso de câncer gástrico. Entretanto quando realizamos a imunomarcação desta proteína, a positividade de *TULP3* não apresentou associação a qualquer histo-tipo, devido ao baixo número de amostras com coloração positiva. Ao analisarmos o perfil de expressão gênica e proteica de *TULP3* em lesões gástricas pré-malignas como gastrite crônica e metaplasia intestinal, identificamos um aumento de expressão gênica no subtipo incompleto de metaplasia. Ainda, ao verificarmos o envolvimento proteico de *TULP3* nos tecidos gástricos, foi identificada uma associação entre a positividade da imunomarcação da proteína com as amostras de metaplasia. Para o câncer colorretal apenas a expressão gênica de *TULP3* foi associada a preditores prognósticos como invasão vascular para o adenocarcinoma retal e invasão linfática para o adenocarcinoma de cólon. Esses achados preliminares indicam um possível papel de *TULP3* na carcinogênese de tumores gástricos e colorretais, entretanto mais estudos são necessários para a definição de *TULP3* como biomarcador para esses tipos tumorais.

Abstract

Gastrointestinal tumors are globally recognized as heterogeneous and aggressive diseases that present bad prognosis for patients who did not undergo resection surgery and the resected ones. The presence of genetic alteration identified in different malignancies, or the expression of specific mRNA and / or tumor proteins, combined with the clinical information, can confer important sources of information for the correct diagnosis, prognosis or prediction of the response of treatments. In particular, mRNA plays a key role in tumor progression and maintenance, and it has a greater potential to reflect cellular phenotypes when compared to changes in DNA. The hypothesis of this study was based on previous findings, in which the *TULP3* transcription factor was identified as a regular master of pancreatic carcinogenesis. In this study the high levels of *TULP3* gene expression were significantly associated with an unfavorable clinical outcome. Furthermore, these increased values of gene expression had no prognostic value in other tumor types, such as breast, ovarian and lung cancers, indicating specificity as a prognostic marker in pancreatic ductal adenocarcinoma. The pancreas, as well as other organs of the gastrointestinal tract, especially the stomach and intestines, have the same embryonic origin, the endoderm. Since that, our hypothesis is that the *TULP3* transcription factor may present a function as a biomarker in gastrointestinal malignancies. This study aimed to investigate the *TULP3* expression profile in gastrointestinal diseases, using data available from gene expression public databases and tissue samples from gastrointestinal tract tumors, especially pancreatic, gastric and colorectal tumors. This work associated, for the first time, the expression of *TULP3* to gastric carcinogenesis. From bioinformatics analyses, using gene expression data available in public databases, the elevated *TULP3* levels were associated with a bad prognosis for the diffuse type of gastric cancer. However, when we performed the immunostaining of this protein, the positivity of *TULP3* was not associated with any histotype, due to the low number of samples with positive staining. When analyzing the gene

and protein expression profile of TULP3 in pre-malignant gastric lesions such as chronic gastritis and intestinal metaplasia, we identified an increase of gene expression in the incomplete subtype of metaplasia. Moreover, when we verified the protein involvement of TULP3 in the gastric tissues, we identified an association between the positivity of TULP3 immunostaining with the metaplasia samples. For colorectal cancer, only the gene expression was associated with prognostic predictors such as vascular invasion for rectal adenocarcinoma and lymphatic invasion for colon adenocarcinoma. These preliminary findings indicate a possible role of TULP3 in the carcinogenesis of gastric and colorectal tumors, however, further studies are needed to define TULP3 as a biomarker for these tumor types.

Capítulo I

Introdução Geral

1. Câncer e biomarcadores

A Agência Internacional para Pesquisa em Câncer (do inglês, *International Agency for Research on Cancer - IARC*) relatou que o impacto global da incidência de câncer mais que dobrou nos últimos 30 anos (Globocan, 2012). Estimativas globais para o ano de 2018 apontam que mais de 17 milhões de novos casos de câncer serão diagnosticados e cerca de 10 milhões de mortes serão relacionadas a neoplasias (Globocan, 2012). No Brasil, o Instituto Nacional do Câncer (INCa) estimou, para o ano de 2018, a ocorrência de 420 mil novos casos de câncer.

Estimated number of new cases in 2018, worldwide, all cancers excl. NMSC, both sexes, all ages

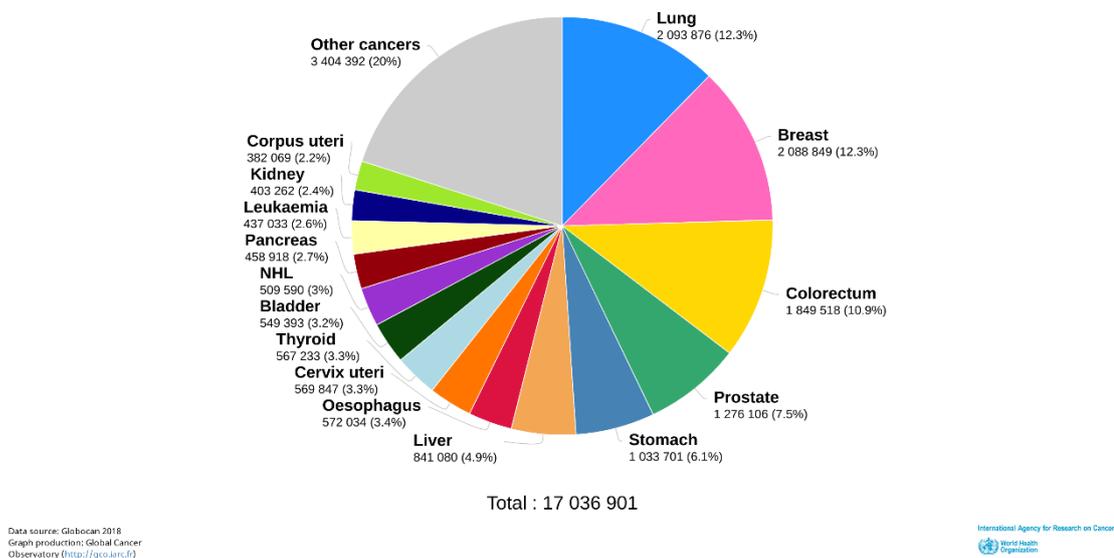


Figura 1. Estimativas mundiais de novos casos para o ano de 2018. Os sítios tumorais que apresentam maiores taxas são demonstrados no gráfico (Globocan, acesso em 28 de Janeiro de 2020).

Estimated number of deaths in 2018, worldwide, all cancers excl. NMSC, both sexes, all ages

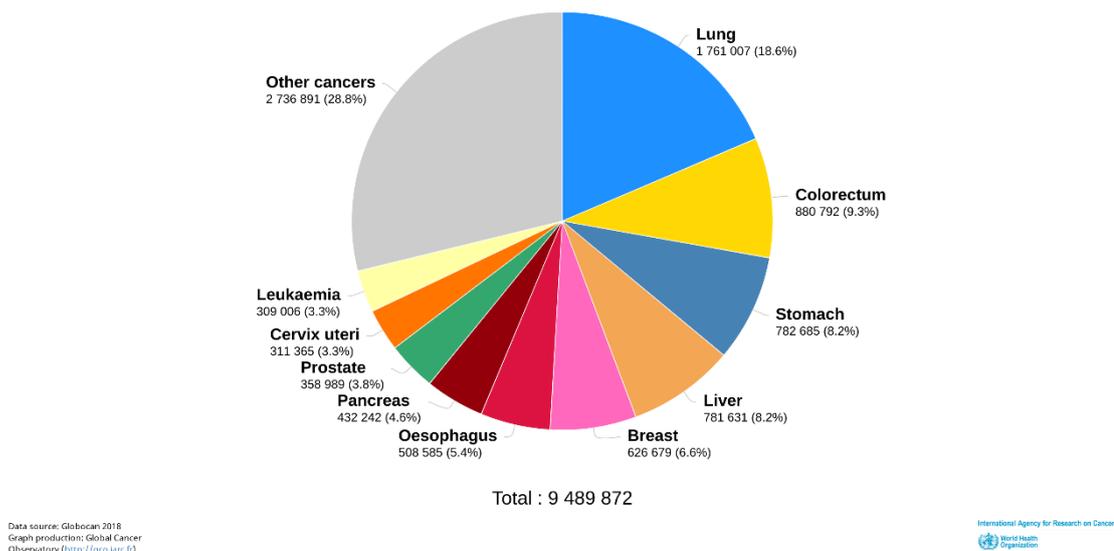


Figura 2. Estimativas mundiais de mortalidade para o ano de 2018. Os sítios tumorais que apresentam maiores taxas de mortalidade são demonstrados no gráfico (Globocan, acesso em 28 de Janeiro de 2020).

Dentre os tipos de câncer, os tumores do trato gastrointestinal (GI) são neoplasias que afetam o aparelho digestivo como um todo, entretanto, acometem, mais especificamente, órgãos como: esôfago, pâncreas, estômago e intestinos (delgado, cólon e reto). Em estudo recente (Costa *et al.*, 2018), foi estimado que, somente as neoplasias GI, representariam cerca de 20% do total de novos casos de câncer diagnosticados no mundo todo e, cerca de 15% das mortes relacionadas por esta doença.

1.1 Câncer gástrico

O câncer gástrico (CG) é reconhecido como a quinta neoplasia mais diagnosticada dentre todos os tipos tumorais, e a terceira neoplasia responsável por mortes por câncer,

entre homens e mulheres (Globocan, 2012). No Brasil, a estimativa da incidência de CG, para o ano de 2018, alcança o terceiro lugar entre homens e o quinto lugar entre as mulheres (INCa).

Anatomicamente, o estômago tem início na junção gastroesofágica e estende-se até o piloro. A parte proximal, que faz ligação com o esôfago, é denominada cárdia, em seguida encontra-se o fundo e o corpo e, por fim, o antro, porção distal do estômago, fazendo comunicação com o duodeno. O estômago é constituído por quatro camadas, denominadas mucosa, submucosa, muscular e serosa (Amin *et al.*, 2017; Sobin & Wittekind, 2004).

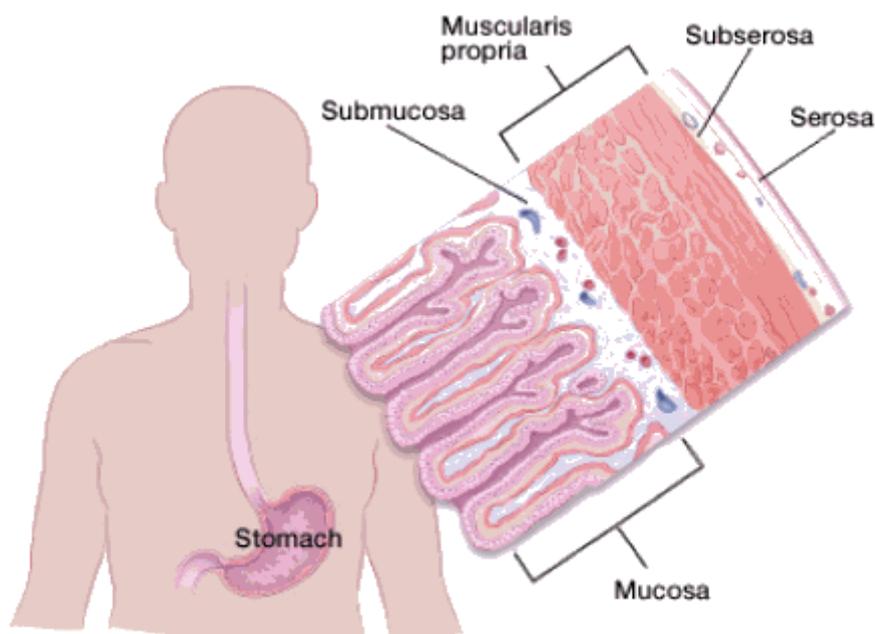


Figura 3. Camadas histológicas que compõe o estômago (American Cancer Society, acesso em 31 de Janeiro de 2019).

O CG pode iniciar-se em qualquer parte do estômago e atingir diferentes camadas de tecidos. O tipo mais comum de CG é o adenocarcinoma, correspondendo a

aproximadamente 95% das neoplasias que acometem esse órgão, e atinge principalmente a camada mucosa (McLean & El-Omar, 2014; Shang & Pena, 2005). O CG pode ser classificado em dois tipos: intestinal e difuso, conforme os critérios histológicos de Laurén (Laurén, 1965). Ambos os histo-tipos de CG estão associados à infecção pela bactéria *Helicobacter pylori* (*H. pylori*). Histologicamente, os adenocarcinomas intestinal e difuso diferem um do outro. Enquanto que ICG é caracterizado pela presença de células tumorais coesivas as quais formam estruturas glandulares, o DCG é composto por células neoplásicas pouco coesas sem formação glandular e mucinoso (Cutsem *et al.*, 2016). A falta de coesão entre as células tumorais do DCG pode ser devido à disfunção na adesão celular, atribuída ao gene *CDH1* (Liu *et al.*, 2013).

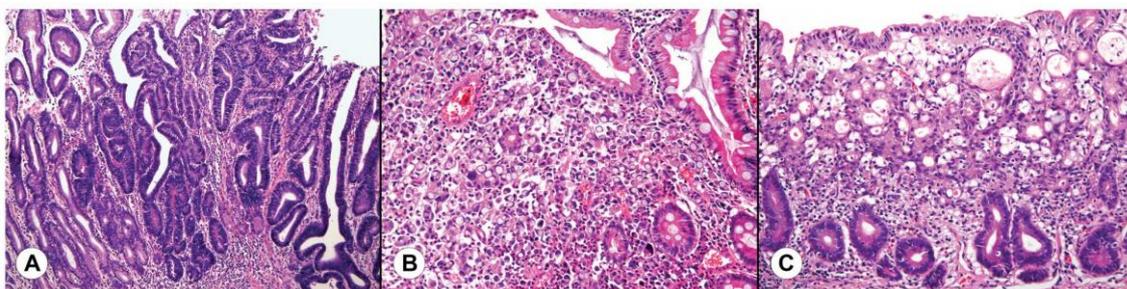


Figura 4. Classificação histológica do adenocarcinoma gástrico, segundo Laurén. (A) tipo intestinal, (B) tipo difuso e (C) tipo misto. (Chong *et al.*, 2014)

O tipo intestinal de CG (ICG) geralmente progride ao adenocarcinoma invasivo através do surgimento das lesões pré-neoplásicas: gastrite atrófica e metaplasia intestinal (Baniak *et al.*, 2016), enquanto que para o tipo difuso de CG (DCG) não há lesões associadas. O DCG é, geralmente, diagnosticado em pacientes jovens os quais apresentam mutações germinativas herdadas no gene supressor de tumor *CDH1* (Hu *et al.*, 2012), entretanto, em trabalho recente de (Cho *et al.*, 2017) foi relatado que

mutações somáticas em *CDH1* também estavam relacionadas ao desenvolvimento precoce de DGC.

A fim de melhorar o entendimento sobre o CG, novas tentativas de classificação dos subtipos do adenocarcinoma gástrico foram realizadas. Em 2010, a Organização Mundial da Saúde (OMS) propôs uma nova classificação baseada, também, em padrões histológicos, agrupando carcinomas gástricos em quatro subtipos: tubular, papilar, mucinoso e pouco coesivo (Fenoglio-Preiser *et al.*, 2000). O subtipo pouco coesivo foi enriquecido com casos de DGC. Já o consórcio The Cancer Genome Atlas (TCGA) realizou uma abordagem genômica e propôs uma classificação molecular do CG dividindo essas neoplasias gástricas em quatro grupos: (1) casos positivos para o vírus Epstein-Barr (EBV), (2) presença de instabilidade microssatélite (MSI), (3) casos genomicamente estáveis (GS) e (4) presença de instabilidade cromossômica (CIN). A maioria dos casos DGC foi agrupada no subtipo GS desta nova classificação (Cancer Genome Atlas Research Network, 2014).

Por fim, o Asian Cancer Research Group (ACRG) classificou o CG com base nos dados da expressão gênica em quatro subtipos: (1) MSI; (2) presença de estabilidade microssatélite e transição epitélio-mesenquimal (MSS/EMT); (3) presença de estabilidade microssatélite com proteína 53 ativa no tumor (MSS/TP53+); e (4) presença de estabilidade microssatélite com proteína 53 inativa no tumor (MSS/TP53-). Todos os subtipos foram associados com resultados clínicos distintos e o subtipo MSS/EMT (enriquecido principalmente com DGC) apresentou o pior prognóstico (Cristescu *et al.*, 2015). Mais recentemente, Ge *et al* (2018) realizaram uma análise de abordagem proteômica do DGC e classificaram o carcinoma difuso em três subtipos (PX1-3), apontando uma distinção entre enriquecimento de vias de sinalização e desfechos clínicos.

O CG é considerado precoce quanto restrito às camadas mucosa e submucosa, independentemente de sua extensão em superfície e da presença ou não de metástases ganglionares, enquanto que os tumores avançados apresentam envolvimento das camadas musculares e serosas, além da mucosa e submucosa (Kim *et al.*, 2011; Tam *et al.*, 2011). O elevado número de casos de CG torna-se um importante problema de saúde pública, uma vez que a maioria dos pacientes é diagnosticada em estágios avançados da doença, quando as opções de tratamento são restritas. Aproximadamente 80% dos pacientes diagnosticados com CG encontram-se em uma fase avançada da doença. Isso se deve, principalmente, pelo fato do CG apresenta-se de maneira assintomática ou pela presença de sintomas pouco alarmantes, como perda de peso, falta de apetite, anemia, dor epigástrica, náuseas e vômito (Schmidt *et al.*, 2005; Bornschein *et al.*, 2010; Correa, 2013).

O prognóstico do CG baseia-se no estadiamento tumoral e está relacionado com a profundidade de invasão neoplásica e com a presença de metástases linfonodais. A abordagem cirúrgica como a gastrectomia e linfadenectomia é considerada a única intervenção com chances de cura aos pacientes diagnosticados com CG (Takahashi *et al.*, 2013; Dikken *et al.*, 2012), uma vez que a quimio e radioterapia apresentam pouco benefício na sobrevivência dos pacientes (INCa; Nagini, 2012).

1.2 Câncer colorretal

O câncer colorretal (CCR) abrange tumores que acometem o cólon (segmento do intestino grosso) e o reto. O CCR é reconhecido, globalmente, como o terceiro tipo de câncer mais comumente diagnosticado em homens (746.000 casos, 10,0% do total de casos relacionados ao câncer) e o segundo em mulheres (614.000 casos, 9,2% do total de casos relacionados ao câncer). Apesar de a mortalidade ser considerada baixa (694.000 mortes, 8,5% do total de mortes relacionadas ao câncer), regiões menos

desenvolvidas apresentam maiores taxas de mortalidade, assim como sobrevida desfavorável para os indivíduos diagnosticados com CCR (Globocan, 2012).

No Brasil, estima-se que cerca de 36 mil pessoas serão diagnosticadas com CCR no ano de 2018, sendo a maioria mulheres (17 mil homens e quase 19 mil mulheres) (INCa). O CCR é caracterizado como uma doença heterogênea, causada pela interação entre fatores genéticos e ambientais. A maioria dos casos de CCR (cerca de 75%) não apresentam componente hereditário, sendo considerados tumores esporádicos, enquanto que 25% dos pacientes diagnosticados apresentam histórico familiar da doença. Fatores ambientais como obesidade, tabagismo, alto consumo de carne vermelha, consumo excessivo de álcool e sedentarismo são associados ao desenvolvimento de CCR. Outros fatores como idade avançada (acima de 60 anos), presença de doença inflamatória intestinal, adenomas ou CCR, histórico familiar de CCR ou síndromes hereditárias como polipose adenomatosa familiar (FAP) e síndrome de Lynch podem aumentar o risco de desenvolvimento do CCR (Labianca *et al.*, 2013; Glimelius *et al.*, 2013; Balmana *et al.*, 2013; DeVita & Rosenberg, 2014).

A sintomatologia varia de acordo com o local de acometimento tumoral. Inicialmente ao estabelecimento do CCR, os pacientes não apresentam sintomas marcantes, entretanto podem exibir diarreia e dor vaga no abdômen (síndrome dispéptica). Nos estágios mais avançados pode-se observar obstrução intestinal com parada de eliminação de fezes e gases associado a vômitos ou mesmo como uma perfuração causando um quadro grave com forte dor abdominal. Um sintoma bastante comum é a perda de sangue em pequena quantidade por períodos prolongados (sangue oculto nas fezes), causando anemia nos pacientes acometidos por CCR (Ballester *et al.*, 2016). A maioria das neoplasias colorretais são classificadas histologicamente como adenocarcinomas, podendo surgir a partir de pólipos adenomatosos inicialmente benignos (Fraum *et al.*, 2016; Kovacheva & Rajpoot, 2016). Da mesma forma que o CG,

o CCR pode atingir diferentes camadas de tecidos, dentre elas, a mucosa, submucosa, muscular e serosa.

A caracterização molecular do CCR foi descrita originalmente por Fearon & Vogelstein (1990), a partir de mutações observadas desde pólipos até o adenocarcinoma. As principais alterações encontradas envolvem os oncogenes *KRAS*, *NRAS*, *BRAF* e *PI3K*, e os supressores tumorais *APC*, *TP53* e *PTEN*. O acúmulo dessas mutações favorece a desregulação de importantes vias de sinalização envolvidas na proliferação, diferenciação e apoptose celular, incluindo as vias de sinalização Wnt/ β -catenina, TGF- β , MAPK e PI3K (Payssonnaux & Eychene, 2001; Calistri *et al.*, 2005; Al-Shamsi *et al.*, 2016). Em 2012, foi publicada uma análise genômica integrativa do CCR a partir de amostras pareadas de tumores e tecidos normais de adenocarcinomas de cólon e reto, identificando a desregulação da sinalização Wnt em ambas as neoplasias (Cancer Genome Atlas Research Network, 2012).

Além do uso de dados genômicos para caracterizar tumores colorretais, os dados transcriptômicos são muito utilizados pois apresentam maior potencial de refletir os fenótipos celulares (Guinney *et al.*, 2015). Para tanto, foi criado o Consórcio para Subtipagem do CCR (CRCSC), que tem como objetivo avaliar a presença ou ausência de padrões principais dos subtipos, através do uso de algoritmos baseados na expressão gênica de tumores colorretais, identificados por seis grupos (A-F), demonstrados na figura 5 (Dienstmann *et al.*, 2014). Considerando que as classificações prévias (Muzny *et al.*, 2012; Perez-Villamil *et al.*, 2012; Schlicker *et al.*, 2012; Budinska *et al.*, 2013; De Sousa E Melo *et al.*, 2013; Marisa *et al.*, 2013; Sadanandam *et al.*, 2013; Roepman *et al.*, 2014), baseadas apenas em dados de expressão gênica, não permitiam a identificação de outros subtipos moleculares que não os principais (subtipo enriquecido com presença de instabilidade microssatélite, e subtipo com alta expressão de genes mesenquimais), Guinney *et al.* (2015) realizaram uma abordagem integrativa multi-ômica para desenvolver uma classificação de subtipos moleculares consenso

(CMS), utilizando dados de expressão gênica associados a dados genômicos (mutações, variação do número de cópias), metiloma, microRNA e dados proteômicos. Foram identificados quatro subtipos moleculares consenso (CMS): CMS1 (MSI Imune, 14%), que apresentavam hipermutação, instabilidade microssatélite e forte ativação imune como características; CMS2 (Canônico, 37%), epitelial, instabilidade cromossômica, marcada ativação das vias de sinalização WNT e MYC; CMS3 (Metabólico, 13%), epitelial, desregulação metabólica evidente; e CMS4 (Mesenquimal, 23%), ativação marcada do fator de crescimento transformador β , invasão estromal e angiogênese.

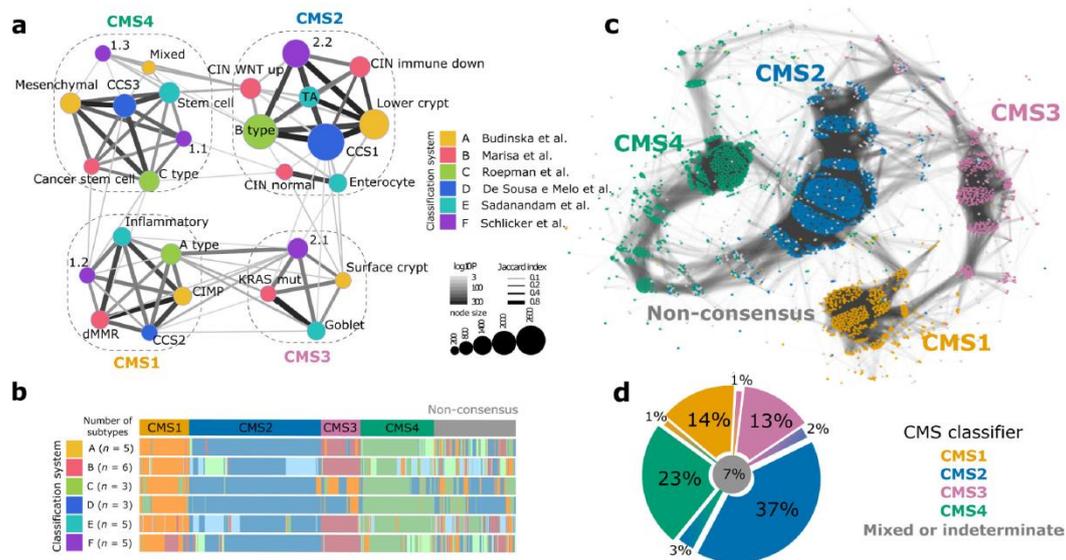


Figura 5. Classificação dos subtipos moleculares consenso de tumores colorretais (CMS). (a) Rede dos subtipos CMS1-4, relacionada aos seis subtipos (A-F) previamente identificados pelo Consórcio para subtipagem do CCR (CRCSC). (b) Distribuição por amostra de cada uma das classificações (A-F) agrupados pelos clusters dos subtipos CMS. As cores dentro de cada linha representam subtipos diferentes. (c) Rede de pacientes diagnosticados com CCR. (d)

Distribuição final dos grupos CMS1–4 (cores sólidas), amostras 'mistas' (cores em gradiente), e amostras indeterminadas (cor cinza).

Apesar dos métodos de rastreio precoce dessa doença, a grande maioria dos casos são diagnosticados em estágios avançados da doença, quando o procedimento cirúrgico já não é mais uma opção, sendo a radio e quimioterapia aplicadas como tratamentos paliativos. A indicação do tratamento quimioterápico é realizada de acordo com o estágio do CCR, para os estágios iniciais da doença, os tratamentos mais comuns envolvem os esquemas 5-FU associado ao leucovorin, e FOLFOX (5-FU, leucovorin e oxaliplatina) (Lin *et al.*, 2017).

1.3 Câncer pancreático

Dentre os tumores pancreáticos, cerca de 85% são classificados como adenocarcinomas e atingem a parte exócrina do pâncreas (Bond-Smith *et al.*, 2012; Ryan *et al.*, 2014). O adenocarcinoma ductal pancreático (ADP) é reconhecido como um tipo de câncer altamente agressivo. Dados globais conferem ao ADP o 7º lugar nas taxas de mortalidade, sendo responsável por cerca de 330.000 mortes relacionadas ao câncer em ambos os sexos. Ainda, são estimados quase 13.000 novos casos de câncer pancreático para o ano de 2020 no Brasil (Globocan, 2012). Dados do INCa, apontam o ADP como responsável por cerca de 2% do diagnóstico dentre todos os tipos de câncer e por 4% do total de mortes por essa doença (INCa).

O diagnóstico precoce de ADP continua sendo um desafio da medicina, uma vez que nos estágios iniciais a doença é assintomática e, geralmente, não apresenta manifestações clínicas até atingir estágios avançados, nos quais os pacientes apresentam dores abdominais não específicas, icterícia, perda de peso e náusea. O

sinal de Courvoisier, caracterizado por apresentar vesícula biliar palpável acompanhada de icterícia sem dores abdominais, é um indicativo de ADP, no entanto, ocorre em menos de 25% dos pacientes (Bond-Smith *et al.*, 2012). Praticamente metade dos casos de ADP são diagnosticados em atendimentos de emergência, nos quais os pacientes exibem sintomas como dores abdominais não específicas, icterícia ou ambos. Quando diagnosticado, cerca de 80-85% dos pacientes apresentam doença irresssecável (i.e., tumores localmente avançados e/ou metástases à distância) (Bilimoria *et al.*, 2007; Vincent *et al.*, 2011; Siegel *et al.*, 2012). O ADP apresenta um prognóstico desfavorável, tanto em pacientes que não sofreram cirurgia de ressecção (sobrevida de 3 a 5% em 5 anos) quanto àqueles que passaram pelo procedimento de ressecção do órgão (sobrevida de 10 a 20% em 5 anos) (Hezel *et al.*, 2006).

A escolha para o tratamento do câncer pancreático é dependente do estágio em que a doença se apresenta. Tumores primários limitados ao pâncreas (estágio I) e tumores localmente invasivos (estágio II) geralmente são ressecáveis enquanto que os tumores localmente avançados (estágio III) e metastáticos (estágio IV) não podem ser removidos cirurgicamente (Bilimoria *et al.*, 2007). As terapias convencionais como rádio e quimioterapia possuem efeitos paliativos para os estágios mais avançados, tornando a cirurgia o único tratamento com chances de cura para o paciente (Yokoyama *et al.*, 2009). O procedimento mais comum é a pancreatoduodenectomia ou procedimento de Whipple, o qual envolve a ressecção do pâncreas proximal, juntamente com o estômago distal, duodeno, ducto biliar distal e vesícula biliar (Bond-Smith *et al.*, 2012). Embora o aperfeiçoamento das técnicas cirúrgicas propicie a ressecção tumoral completa, a maioria dos pacientes submetidos ao procedimento apresenta recorrência. A fim de aumentar as chances de cura para os pacientes submetidos à cirurgia de ressecção são aplicados quimioterápicos como gemcitabina, cisplatina e 5-fluoro-uracil (5-FU) sozinhos ou em combinação, como terapia sistêmica adjuvante. As terapias adjuvantes, caracterizadas pela combinação de químico e radioterapia, são indicadas, em alguns

casos, antes do procedimento cirúrgico, assim como aos pacientes que não são candidatos à cirurgia (Herrerros-Villanueva *et al.*, 2012). Entretanto, muitos pacientes apresentam resistência à maioria das formas de tratamento desenvolvidas até hoje, o que somadas, tornam o ADP uma doença tão letal (Li *et al.*, 2004).

O ADP é uma doença associada a idades avançadas, apresentando maior incidência entre 70 e 80 anos, quando atinge um risco aumentado de desenvolvimento da doença em 40 vezes (Bardeesy & DePinho, 2002). Fatores ambientais, como o tabagismo, podem contribuir para o desenvolvimento do câncer pancreático, uma vez que este hábito aumenta as taxas de risco de 1.6 para 5.4 (Li *et al.*, 2004). O consumo excessivo de álcool, associado à pancreatite crônica, também pode conduzir ao estabelecimento do câncer (Nitsche *et al.*, 2011). Outros fatores de risco como síndromes genéticas que podem aumentar a probabilidade do desenvolvimento do ADP incluem: pancreatite hereditária, fibrose cística, síndrome de Peutz-Jeghers, síndrome de Lynch (também conhecida como câncer colorretal hereditário não-polipóide, HNPCC), síndrome do nevo atípico e melanoma familiar (também conhecido como síndrome familiar atípica de múltiplos nevos melanocíticos, FAMMM), câncer de mama e ovário hereditários e câncer pancreático hereditário (Decker *et al.*, 2010).

O ADP é uma neoplasia sólida de caráter infiltrante o qual pode surgir a partir de precursores denominados neoplasias intraepiteliais pancreáticas (PanIN), neoplasias intraductais papilíferas mucinosas (IPMN) e as neoplasias císticas mucinosas (MCN). As lesões IPMN e MCN são caracterizadas pela produção de mucina, sendo visíveis em exames de imagem (Maitra & Hruban, 2008; Bosman *et al.*, 2010). Essas lesões precursoras, em especial as PanINs, estabeleceram um modelo morfológico e genético de progressão tumoral que incorpora alterações no genoma. A progressão histológica de neoplasias intraepiteliais de baixo grau (PanIN-1) para intermediário (PanIN-2) e alto (PanIN-3) graus e, posteriormente para o carcinoma invasivo está associada ao acúmulo de alterações (Hruban *et al.*, 2001). As alterações sequenciais observadas a

partir de células pancreáticas normais até o adenocarcinoma, descritas até o presente momento, são encontradas nos genes *KRAS*¹, *CDKN2*, *TP53*, *SMAD4/DPC4* e *BRCA2* (Hruban *et al.*, 2001). A identificação de mutações ocorre de forma diretamente proporcional com a variação do grau histológico de atipia, ou seja, quanto maior o grau de atipia, maior o número de mutações genéticas encontradas (Ardengh *et al.*, 2008).

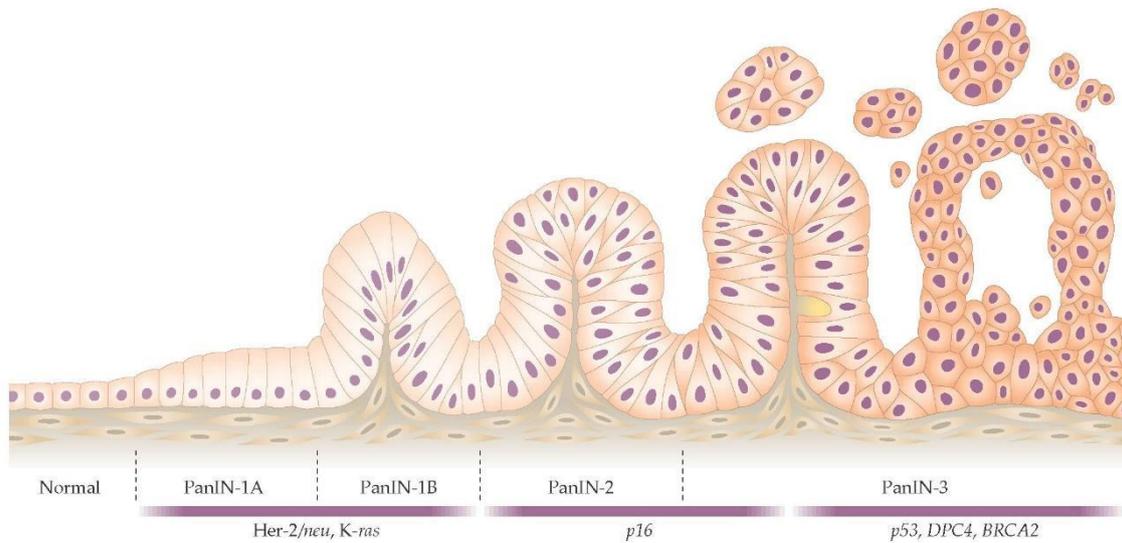


Figura 6. Modelo de progressão do câncer pancreático. Alterações genéticas são acumuladas durante a transformação do ducto normal em lesões intraepiteliais ao câncer invasivo (Hruban *et al.*, 2001).

1.4 Biomarcadores tumorais

Os pacientes com neoplasias GIs frequentemente apresentam sintomas inespecíficos no início da doença ou mesmo assintomáticos em estágios avançados, resultando no diagnóstico tardio quando poucas opções terapêuticas estão disponíveis. A dificuldade de diagnóstico é resultado não só da inespecificidade dos sintomas, mas

¹Genes e RNA mensageiro (RNAm) de *Homo sapiens* são apresentados com todas as letras maiúsculas e formatação em itálico.

também das limitações metodológicas relacionadas. Apesar de existir exames de imagem para a identificação de neoplasias GIs como ultra-sonografia (e.g.: ecoendoscopia), tomografia computadorizada e ressonância magnética, essas técnicas são ineficazes para a detecção dos tumores precocemente, já que não são confiáveis para a identificação de tumores menores do que 1 a 2 cm (Ardengh *et al.*, 2008). Além disso, os exames de imagem não são capazes de distinguir com precisão os tumores de lesões pré-malignas no mesmo tecido. Ainda, em alguns casos de doença avançada ou resultados inconclusivos por meio de exames de imagem e exames séricos, os pacientes são submetidos a procedimentos invasivos, como laparotomia exploratória, para possibilitar a determinação do estadiamento e as decisões terapêuticas futuras (Hidalgo, 2010; Mayo *et al.*, 2009).

Apesar de existir biomarcadores detectáveis no soro dos pacientes com suspeita de neoplasias GI, estes são, muitas vezes inespecíficos, sendo encontrados em mais de um tecido, o que dificulta o correto diagnóstico da doença. O método ouro para diagnóstico definitivo continua sendo a análise histopatológica, aliada a métodos de imunohistoquímica (IHQ) para identificação dos tipos tumorais/lesões pré-malignas e da conduta médica. A inespecificidade dos marcadores séricos e dos métodos de imagem para o diagnóstico definitivo demanda estudos que utilizem amostras de tecidos tumoral e não-tumoral adjacente a fim de identificar molecularmente as diferenças existentes entre estes tecidos e também no estadiamento dos tumores.

Devido a estas dificuldades, torna-se urgente a necessidade de identificação de novos alvos moleculares capazes de: detectar precocemente as neoplasias GIs (biomarcadores de diagnóstico); prever ou monitorar a resposta dos pacientes em relação à doença diagnosticada ou ainda à resposta ao tratamento (biomarcadores de prognósticos); e auxiliar na definição de intervenções terapêuticas (biomarcadores preditivos). Neste sentido, a busca por novos alvos tem encontrado resultados promissores nos campos da genômica, transcriptômica e proteômica. (Gonzaga *et al.*, 2017).

Os marcadores diagnósticos são classificados como aqueles que permitem a identificação da fase ativa da doença, bem como a identificação de pacientes em risco. Os marcadores prognósticos podem ser definidos como aqueles que estão associados ao desfecho clínico, seja ele o risco de morte devido ao câncer ou ao risco de recidiva da doença, independentemente de qualquer tratamento ou intervenção. Entende-se como fator prognóstico qualquer marcador, à época do diagnóstico ou do tratamento cirúrgico, que possa ser associado ao tempo livre de doença ou sobrevida global na ausência de terapia sistêmica adjuvante. Esses fatores podem ser usados para prever a história natural do tumor. O fator prognóstico irá selecionar as pacientes, que, provavelmente, terão recorrência sem terapia sistêmica e, por isso, se beneficiarão da implementação dessa estratégia, apesar da morbidade associada ao seu recurso. A aplicação criteriosa de fatores prognósticos estabelecidos e validados é essencial para se aumentar a utilização de terapia sistêmica e, conseqüentemente, diminuir a mortalidade (Ludwig & Weinstein, 2005; Payne *et al.*, 2008).

O fator preditivo representa qualquer marcador associado a uma resposta em uma terapia específica, ou seja, ele seleciona a melhor terapia para cada paciente (Bhargava *et al.*, 2008). Algumas variáveis podem ser, ao mesmo tempo, prognósticas e preditivas, o que confere maior complexidade à análise desses fatores. A seguir, são listados alguns exemplos de fatores preditivos e prognósticos importantes: idade, tamanho tumoral, subtipo histológico de neoplasias, grau de diferenciação tumoral.

No geral, os biomarcadores moleculares podem ser classificados como marcadores baseados em DNA, RNA ou proteínas. Os marcadores baseados em DNA incluem polimorfismos de base única (SNPs), aberrações cromossômicas, alteração do número de cópias, instabilidade microssatélite, e metilação diferencial na região promotora. Os marcadores baseados em RNA incluem a alta ou baixa expressão dos transcritos de RNA mensageiro, e também os RNAs reguladores (e.g.: microRNAs). Os marcadores proteicos incluem: receptores de superfície, tumores antigênicos, estados de

fosforilação, determinantes carboidratos, e peptídeos liberados por tumores no soro, urina, e outros fluidos corporais (Brooks, 2012).

A combinação de técnicas de alto rendimento incluindo sequenciamento de DNA, e quantificação de RNA e proteínas podem ser utilizadas em conjunto para a identificação de biomarcadores candidatos capazes de diferenciar diferentes tipos tumorais (Brooks, 2012, Nalejska *et al.*, 2014). Ainda, a detecção destes marcadores moleculares em amostras de tecido doente pode prover informações importantes em diversos aspectos do manejo clínico dos pacientes, como no diagnóstico (preferencialmente precoce, detectando lesões clinicamente silenciosas), na definição dos diferentes subtipos da doença; na progressão tumoral, no prognóstico, ou ainda no monitoramento da resposta ao tratamento (Brooks, 2012).

A maioria dos tumores GIs apresentam acúmulo de alterações no genoma, transcriptoma e/ou proteoma, o que pode levar à perda do controle dos mecanismos regulatórios, causando proliferação desenfreada e o avanço da doença (Sharma *et al.*, 2018).

Tumores GIs são descritos como altamente heterogêneos, nos quais os pacientes diagnosticados exibem perfis molecular e genético distintos uns dos outros, mesmo tratando-se do mesmo tipo tumoral. Apesar de muitos estudos serem realizados para a identificação de biomarcadores específicos para estas neoplasias, a maioria dos marcadores moleculares estudados não apresenta resultados promissores na etapa de validação (Matsuoka & Yashiro, 2018).

1.4.1 Biomarcadores séricos e moleculares para o CG

Para o CG, os biomarcadores séricos comumente utilizados e validados na clínica são o antígeno carcinoembrionário (CEA) e o antígeno carboidrato 19-9 (CA19-9). O CEA é uma glicoproteína de membrana envolvida na adesão celular. Uma vez que altos

níveis séricos de CEA são encontrados em pacientes com estágios avançados de CG, este marcador não é considerado eficaz em métodos de screening, e sim, como preditor de fator de risco de metástase hepática independente da recidiva (Matsuoka & Yashiro, 2018). O CA19-9 é um antígeno glicolipídico expresso na superfície de células endoteliais, este marcador é comumente utilizado em pacientes com suspeita de neoplasias GIs, entretanto, níveis elevados de CA19-9 são encontrados em diversos tipos tumorais, em especial no CG, ADP e CCR (Shimada *et al.*, 2014).

Neste sentido, a busca por marcadores moleculares faz-se necessária para o melhor manejo clínico dos pacientes diagnosticados com CG. Devido à crescente disponibilidade de informações provenientes de análises ômicas e ao melhor entendimento da biologia molecular desta neoplasia, muitos estudos tem demonstrado o importante papel de HER2 no estabelecimento do CG, sendo este o primeiro marcador molecular a ser utilizado na prática clínica (Boku, 2014).

HER2² é uma proteína de membrana celular, classificada como um proto-oncogene e codificada pelo gene *ERBB2* no cromossomo 17, a qual liga-se ao receptor tirosina cinase e, quando desregulada, ativa uma série de vias de sinalização envolvidas na proliferação celular e manutenção tumoral (Moasser, 2007). Nos carcinomas gástricos, a amplificação de *HER2* tem sido encontrada de maneira variável de 7 a 34% dos casos (Takehana *et al.*, 2002; Hofmann *et al.*, 2008). Esta variabilidade é decorrente principalmente das diferentes metodologias utilizadas para a pesquisa da expressão e/ou amplificação do HER2 (Allgayer *et al.*, 2000; Takehana *et al.*, 2002; Garcia *et al.*, 2003). A superexpressão de HER2 nos carcinomas gástricos tem sido associada à pior sobrevida dos pacientes (Allgayer *et al.*, 2000; Shinohara *et al.*, 2002; Begnami *et al.*, 2011). Baseados nestes achados, a associação do tratamento quimioterápico (capecitabina ou 5-FU e cisplatina) e Trastuzumabe tem sido apontada como uma

² Proteínas de *Homo sapiens* são apresentadas com todas as letras maiúsculas sem formatação em itálico.

terapia alternativa para os pacientes com câncer gástrico avançado. Dados do estudo clínico Trastuzumabe for Gastric Cancer (ToGA), (Bang *et al.*, 2010) protocolo BO18255, demonstraram que pacientes com carcinomas gástricos com amplificação de *HER2* que fizeram uso deste esquema terapêutico apresentaram significativa redução do tamanho tumoral, diminuição das metástases abdominais e maiores taxas de sobrevida global e livre de doença (Bilous *et al.*, 2010). A positividade para HER2 nos carcinomas gástricos, definida pelo ToGA, foi a expressão imuno-histoquímica classificada como 3+ e/ou amplificação pelo FISH+ (razão HER2/cromossomo 17 > 2,0).

1.4.2 Biomarcadores séricos e moleculares para o CCR

Para o CCR, os marcadores séricos utilizados na prática clínica incluem o CEA e o CA 19-9 (Staab *et al.*, 1985; Yu *et al.*, 2017). Em geral, os níveis de CEA encontram-se elevados nos casos de recorrência tumoral, sendo este parâmetro considerado um marcador de vigilância pós-operatória (Yakabe *et al.*, 2010; aPark *et al.*, 2009; Zeng *et al.*, 1993). Embora a especificidade do CA19-9 na detecção de CCR seja elevada (96%) (Goldberg *et al.*, 1989), sua sensibilidade é muito baixa (de apenas 23%) e sua utilidade como marcador de prognóstico ainda é controversa (bPark, *et al.*, 2009).

Um exemplo de marcador molecular usado para o diagnóstico de CCR são os marcadores de instabilidade microssatélite (MSI), os quais apresentam mutações em genes envolvidos no reparo de sequências de DNA, como os genes *MLH1*, *MSH2*, *MSH6*, *PMS2*. Estes genes mutados do sistema de reparo MMR são traduzidos em proteínas disfuncionais que não são capazes de corrigir os erros ocorridos espontaneamente durante a replicação do DNA, ocorrendo, portanto, o acúmulo de sequências repetidas de DNA, denominadas de regiões microssatélites (Thibodeau *et al.*, 1993; Lindor *et al.*, 2002). A presença de MSI em pacientes diagnosticados com CCR (estágio II) é um indicador de bom prognóstico e também que estes pacientes não

se beneficiam do tratamento quimioterápico adjuvante com 5-FU, quando comparados aos pacientes que realizaram apenas a cirurgia de ressecção. Apesar de ser um bom indicador, apenas de 12-15% dos tumores CCR apresentam MSI (Kawakami, *et al.*, 2015).

Diversos biomarcadores relacionados a DNA, RNAm, proteínas, dentre outros, têm sido descritos em estudos ômicos recentes. A identificação de biomarcadores nas mais variadas neoplasias pode contribuir para o progresso em estudos relacionados à biologia tumoral, seleção de estratégias terapêuticas apropriadas e eficiência em programas de acompanhamento de pacientes diagnosticados com câncer (Matsuoka & Yashiro, 2018).

1.4.3 Biomarcadores séricos e moleculares para o ADP

Para o ADP, o biomarcador sérico comumente utilizado e validado na clínica é o antígeno de superfície CA19-9. Este carboidrato de superfície, entretanto, não é expresso em pacientes negativos para os antígenos 'a' e 'b' do grupo sanguíneo de Lewis, correspondendo a 10% dos pacientes diagnosticados com ADP. Esses pacientes não sintetizam CA19-9 nem mesmo nos estágios avançados da doença. Apesar de inespecífico para o diagnóstico do ADP, este marcador possibilita o monitoramento da doença e detecção precoce de recaída em pacientes já diagnosticados (Harsha *et al.*, 2009). Ainda, os níveis de CA19-9 não são capazes de diferenciar condições não-malignas, como a pancreatite, de tumores pancreáticos (Ballehaninna & Chamberlain, 2013; Khan *et al.*, 2017). Sendo assim, um melhor entendimento da patogênese molecular do ADP pode ter grande utilidade na identificação de novos biomarcadores e estratégias terapêuticas efetivas.

O carcinoma invasivo (ADP) se estabelece através de alterações nos genes *KRAS*, *CDKN2*, *TP53*, *SMAD4/DPC4* e *BRAC2*, a partir de células pancreáticas normais, lesões precursoras até o adenocarcinoma (Hruban *et al.*, 2001). Tanto no ADP quanto nas lesões precursoras (PanIN, IPMN e MCN) foram identificadas mutações no gene *KRAS*, entretanto, como este é um membro da família RAS de GTPases, as quais estão envolvidas em grande variedade de funções e vias de sinalização celulares incluindo proliferação, diferenciação e sobrevivência, torna-se muito complexo prever o comportamento dos tumores que apresentam mutações neste gene (Chang & Kundranda, 2017).

2. Biologia molecular e análise computacional

Em biologia molecular, o fluxo de informação gênica, contida no DNA, avança quando é transcrito em RNA mensageiro (RNAm) e, posteriormente, traduzida em proteína. A correta execução de cada passo é essencial, uma vez que a ocorrência de falhas nesse processo, sem a devida correção por mecanismos regulatórios, afetam todos os passos subsequentes, como exemplo, alterações na expressão gênica que podem resultar na expressão aberrante de proteínas (Fernandes-Zapico *et al.*, 2003).

Muitos autores reportam que a expressão gênica é considerada um marcador essencial para a avaliação diagnóstica e prognóstica em diversos tipos tumorais (Golub *et al.*, 1999; Ntzani & Ioannidis, 2003; Wang *et al.*, 2005; Sotiriou & Piccart, 2007; van't Veer & Bernards, 2008; Mehta *et al.*, 2010; Cristescu *et al.*, 2015; Guinney *et al.*, 2015). Há mais de uma década, a análise da expressão gênica foi proposta como método para complementar os esquemas de classificação baseados na morfologia tumoral, uma vez que tumores do mesmo histo-tipo podem apresentar desfechos clínicos consideravelmente diferentes (Golub *et al.*, 1999; Wan *et al.*, 2015).

É cada vez maior o volume de dados produzido em estudos de biologia molecular e celular através de técnicas como sequenciamento, microarranjo, sequenciamento de DNA complementar seguido de quantificação (RNA-Seq). Essas técnicas são capazes de produzir uma enorme quantidade de dados de maneira rápida a custos cada vez mais acessíveis (Brooks, 2012; Nalejska *et al.*, 2014). A crescente demanda na busca de maior capacidade de armazenamento de dados deu origem a repositórios públicos como o Gene Expression Omnibus (GEO) no ano de 2000 (Barrett *et al.*, 2011). Da mesma forma, outros bancos de dados surgiram, como o The Cancer Genome Atlas (TCGA), que engloba informações multidimensionais sobre alterações genômicas de 33 tipos tumorais. A compilação desse conjunto de dados surge a partir de uma colaboração entre o National Cancer Institute (NCI) e o National Human Genome

Research Institute (NHGRI), que compreende mais de 2 petabytes de dados genômicos, disponibilizados publicamente. Essas informações podem ser utilizadas em estudos para auxiliar a comunidade científica na busca do aprimoramento da prevenção, diagnose e tratamento de neoplasias humanas.

O grande desafio da era ômica baseia-se na busca de informações biologicamente relevantes, o que requer análises robustas que consigam processar as interações complexas, geradas através dessas técnicas de alto rendimento. Neste sentido, o uso de ferramentas de biologia computacional tem encontrado grande utilidade nas análises ômicas.

A expressão de RNAm e/ou proteínas tumorais específicas, ou ainda, alguma alteração genética, identificadas em diferentes neoplasias, aliadas às informações clínicas, podem conferir importantes fontes de informações para o correto diagnóstico, prognóstico ou predição a resposta de tratamentos. Em especial, o RNAm apresenta papel fundamental na progressão e manutenção tumoral e, além disso, tem maior potencial de refletir os fenótipos celulares quando comparados a alterações no DNA (Cristescu *et al.*, 2015).

Em estudo de Sartor *et al* (2014), foram utilizados dados de expressão gênica de pacientes diagnosticados com ADP, provenientes de microarranjos depositados em bancos de dados públicos, como GEO (Edgar *et al.*, 2002) e ArrayExpress (Rustici *et al.*, 2013). A fim de identificar os reguladores mestres da transcrição em ADP, esses dados de expressão gênica foram analisados através dos métodos Algorithm for the Reconstruction of Accurate Cellular Networks (ARACNe) (Margolin *et al.*, 2006) e Master Regulator Analysis (MRA) (Carro *et al.*, 2010) implementados no ambiente R (R Development Core Team, 2011). Como resultado foi ressaltado o possível papel de Tubby-like Protein 3 (*TULP3*) como biomarcador específico para ADP. Neste estudo verificou-se que os elevados níveis de expressão de *TULP3* apresentaram associação significativa com um desfecho clínico desfavorável. Ainda, esses valores aumentados

de expressão gênica não possuíam qualquer valor prognóstico em outros tipos tumorais, como cânceres de mama, ovário e pulmão, indicando a especificidade como marcador prognóstico no ADP. Apenas outro estudo (Harada *et al.*, 2009) descreveu alterações dos níveis transcricionais de *TULP3*, envolvendo tanto lesões intraepiteliais (PanIN) como o carcinoma invasivo (ADP).

O pâncreas, glândula derivada do endoderma, possui a mesma origem embrionária que os órgãos do trato gastrintestinal como o esôfago, estômago e intestinos (Tam *et al.*, 2003; Pinchuk *et al.*, 2010; Fausett & Klingensmith, 2012; Jennings *et al.*, 2015; Kim & Shivdasani, 2016). Sendo esses órgãos de mesma origem embrionária, nossa hipótese é de que o fator de transcrição *TULP3* possa apresentar uma função como biomarcador em neoplasias do sistema gastrintestinal.

3. Tubby-Like Protein 3 (TULP3)

Ao estudar a síndrome da diabetes-obesidade monogênica, Coleman e Eicher (1990) associaram o fenótipo “roliço” (*tubby*) de ratos à síndrome espontânea da obesidade de início tardio, uma vez que a progressão do fenótipo é lenta. Contudo, estudos posteriores demonstraram que estes ratos possuíam deficiências na audição e visão (Ohlemiller *et al.*, 1995). Estas observações possibilitaram a identificação de um novo gene, denominado *Tub*³, que, quando mutado, era responsável pelo fenótipo *tubby* em ratos (Kleyn *et al.*, 1996).

A característica comum a todas as proteínas que constituem esta família é a presença de um domínio *tubby* positivamente carregado na porção carboxi-terminal, o qual possui uma sequência conservada de localização nuclear, permitindo sua atuação como fator de transcrição, sendo também capaz de interagir com fosfoinosítídeos de membrana (Boggon *et al.*, 1999). Enquanto que a região carboxi-terminal é conservada, a região amino-terminal apresenta grandes variações e funções distintas, permitindo aos membros TULP3, TULP2 e TUB de ligar-se com transporte ciliar intraflagelar do complexo A (IFT-A). Uma vez associado ao IFT-A, TULP3 está comprometido ao transporte vesicular dos receptores acoplados à proteína G (GPCR) ao cílio e ao transporte retrógrado e anterógrado no interior do cílio primário (Mukhopadhyay & Jackson, 2011).

TULP3 encontra-se adjunto a PIP₂ (phosphatidylinositol-4,5-bisphosphate) através da associação do domínio *tubby* a este fosfolípídeo de membrana. O desligamento de TULP3 da membrana é promovido a partir da ativação da GPCR, o que gera a hidrólise de PIP₂ e a liberação de TULP3 no citoplasma. O fato de TULP3 não se encontrar associado a PIP₂ ou IFT-A pode resultar na sua translocação nuclear e ativação da

³ Genes e RNAm de ratos e camundongos são apresentados com a primeira letra maiúscula e formação em itálico.

transcrição. Da mesma forma, quando a ligação TULP3-(IFT-A) é prejudicada, TULP3 não entra no cílio e pode ser translocado para o núcleo e, ainda, sua ausência no cílio estimula a sinalização Hedgehog (HH), potencializando a transcrição (Mukhopadhyay *et al.*, 2010).

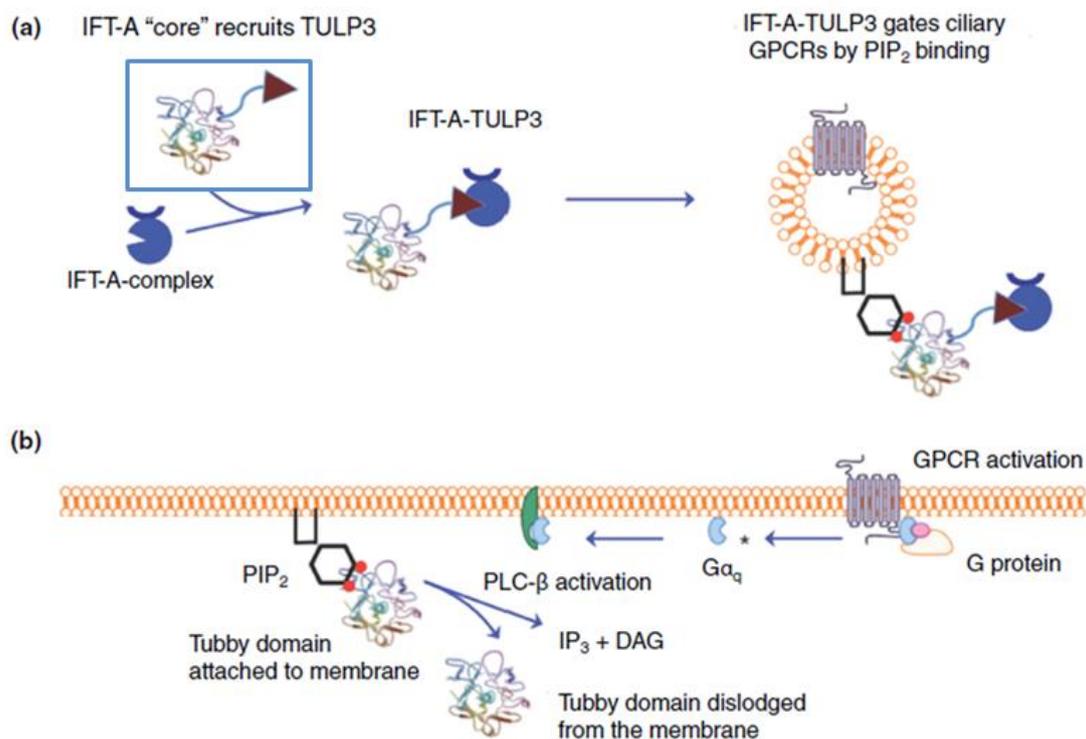


Figura 7. Domínios de ligação de TULP3. (a) TULP3 se associa ao core IFT-A através da região N-terminal. TULP3-(IFT-A) aprisionam GPCRs ciliares à vesícula através da ligação de TULP3 com PIP₂ são transportados via IFT-A de forma retrógrada e anterógrada no cílio primário. (b) Domínio tubby de TULP3 é ancorado à PIP₂, na membrana. A ativação da GPCR causa a hidrólise de PIP₂ e a liberação de TULP3 no citoplasma. (IFT-A) transporte intraflagelar do complexo A. (IP₃) inositol triphosphate. (DAG) diacylglicerol. (PCL) phospholipase C. (Mukhopadhyay & Jackson, 2011, modificado).

O gene *TULP3* está mapeado no braço curto do cromossomo 12 (região 12p13), (Nishina *et al.*, 1998) está anotado no Genome Reference Consortium Human Build 38

(GRCh38), tem 1482bp de extensão (nucleotídeos 2,877,233 - 2,941,140) e está estruturado em 12 éxons.

4. Justificativa

Os tumores gastrintestinais são reconhecidos mundialmente como doenças heterogêneas e agressivas que apresentam um prognóstico desfavorável tanto para os pacientes que não realizaram a cirurgia de ressecção quanto àqueles que passaram pelo procedimento de ressecção do órgão. A hipótese desse estudo baseou-se em achados prévios, nos quais o fator de transcrição *TULP3* foi apontado como regulador mestre da carcinogênese pancreática, e quanto à similaridade embrionária dos tecidos gastrintestinais, em especial do pâncreas, estômago e intestinos.

A disponibilização de dados gerados através de técnicas ômicas para o estudo de neoplasias pode contribuir para o entendimento da biologia tumoral quando associados a dados clínicos, podendo ainda, contribuir para a identificação de informações importantes relacionadas ao diagnóstico, à definição dos diferentes subtipos da doença, à progressão tumoral e ao prognóstico desses pacientes.

5. Objetivos

5.1 Objetivo geral

Investigar o perfil de expressão gênica de *TULP3* em tecidos gastrintestinais, especialmente em lesões e tumores gástricos, colorretais e pancreáticos, utilizando dados disponíveis em bancos de dados públicos como o GEO e o TCGA, e validação experimental através da investigação da expressão proteica.

5.2 Objetivos específicos

Avaliar a expressão de *TULP3* como biomarcador em tumores gástricos e lesões pré-malignas associadas à progressão tumoral utilizando dados disponíveis em bancos de dados públicos, e através da técnica de imunohistoquímica, associando sua expressão ao desfecho e informações clínicas.

Avaliar a expressão gênica de *TULP3* como biomarcador em tumores colorretais utilizando dados disponíveis em bancos de dados públicos como o GEO e o TCGA e associar sua expressão ao desfecho e informações clínicas.

Avaliar a expressão de *TULP3* como biomarcador em amostras de tecidos pancreáticos através da quantificação de RNAm e da imunocoloração, associando sua expressão ao desfecho e informações clínicas.

Capítulo IV

Manuscrito III

Manuscrito publicado no periódico PlosOne em 14 de Janeiro de 2019

RESEARCH ARTICLE

TULP3: A potential biomarker in colorectal cancer?

Ivaine Taís Sauthier Sartor^{1,2*}, Mariana Recamonde-Mendoza^{3,4}, Patricia Ashton-Prolla^{1,2,5}

1 Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, PPGBM-UFRGS, Porto Alegre, Rio Grande do Sul, Brazil, **2** Laboratório de Medicina Genômica, Serviço de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, LMG-SPE-HCPA, Porto Alegre, Rio Grande do Sul, Brazil, **3** Departamento de Informática Aplicada, Instituto de Informática, Universidade Federal do Rio Grande do Sul, INF-UFRGS, Porto Alegre, Rio Grande do Sul, Brazil, **4** Núcleo de Bioinformática, Serviço de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Porto Alegre, Rio Grande do Sul, Brazil, **5** Departamento de Genética, Universidade Federal do Rio Grande do Sul, DEGEN-UFRGS, Porto Alegre, Rio Grande do Sul, Brazil

* ivaine.sauthier@gmail.com



OPEN ACCESS

Citation: Sartor ITS, Recamonde-Mendoza M, Ashton-Prolla P (2019) *TULP3*: A potential biomarker in colorectal cancer? PLoS ONE 14(1): e0210762. <https://doi.org/10.1371/journal.pone.0210762>

Editor: Yves St-Pierre, Institut National de la Recherche Scientifique, CANADA

Received: September 20, 2018

Accepted: January 1, 2019

Published: January 14, 2019

Copyright: © 2019 Sartor et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This study was supported in part by grants from Fundo de Incentivo à Pesquisa, Hospital de Clínicas de Porto Alegre (FIPE-HCPA 16-0032), FAPERGS (16/2551-0000520-6), and I.T.S.S. and P.A-P CNPq scholarships (140045/2015-5 and 307826/2017-1).

Competing interests: The authors have declared that no competing interests exist.

Abstract

Colorectal cancer (CRC) is the second most common cancer in women and the third most common cancer in men globally. The identification of differentially expressed genes associated to patient’s clinical data may represent a useful approach to find important genes in CRC carcinogenesis. Previously, the *TULP3* transcription factor was identified as a possible prognostic biomarker in pancreatic ductal adenocarcinoma. Considering that pancreatic and colorectal tissues have the same embryonic origin, we investigated the profile of *TULP3* expression in CRC hypothesizing that it may have a role in its development. We comparatively analysed *TULP3* gene expression in CRC and normal adjacent colonic tissue and assessed association of expression profiles with survival and clinicopathological information, using publicly available datasets. *TULP3* expression levels were increased in CRC when compared to the adjacent non-tumoral tissue. In addition, higher *TULP3* gene expression was associated to lymphatic and vascular invasion in colon adenocarcinoma (COAD) and rectum adenocarcinoma (READ), respectively. In summary, our results point to a possible role of *TULP3* as a diagnostic and prognostic biomarker in CRC. Additional studies are necessary to confirm these preliminary findings.

Introduction

The global impact of cancer incidence and mortality is increasing, according to the International Agency for Research on Cancer (IARC). In 2020, more than 17 million new cancer cases will be diagnosed and about 10 million deaths in both sexes will be related to cancer [1]. Among all types of cancer, gastrointestinal (GI) tumors (those affecting the esophagus, pancreas, stomach and intestines) are responsible for approximately 20% of all diagnosed cancers worldwide and about 15% of cancer-related deaths [2].

Colorectal cancer (CRC) encompasses tumours affecting the colon and rectum. CRC is globally recognized as the third most commonly diagnosed type of cancer in men (746,000 cases, 10.0% of all cancer cases) and the second in women (614,000 cases, 9.2% of all cancer cases). Although mortality is considered low (694,000 deaths, 8.5% of total cancer deaths), less developed regions present higher mortality rates, as well as unfavorable survival for individuals diagnosed with CRC [1].

mRNA expression is recognized to play an important role in tumour cell survival and tumour progression, and comparative gene expression profiles between tumours and adjacent tissues has been important in the identification of biomarkers. [3]. Thus, in colorectal and adjacent non-tumoral tissues, analysis of differential gene expression, coupled with analysis of patient's clinical data may represent a useful approach to identify important genes in CRC carcinogenesis [3, 4]. In a previous study based on computational analyses, Sartor *et al.* [5], identified the Tubby-like Protein 3 (*TULP3*) transcription factor as a master regulator of carcinogenesis in pancreatic ductal adenocarcinoma (PDAC) from microarray-based expression profiles. Additionally, the authors observed higher *TULP3* expression associated to a poor prognosis in PDAC patients. Given the similarity in embryonic development of pancreatic and intestinal tissues [6–10], we hypothesized that *TULP3* may also play a role in colorectal carcinogenesis.

Methods

Data acquisition and preprocessing

Gene expression data from patient biopsies and control samples were obtained from The Cancer Genome Atlas–TCGA Research Network (<http://cancergenome.nih.gov>) and Gene Expression Omnibus (GEO) [11] databases. From TCGA, we analysed the COAD and READ studies (Colon Adenocarcinoma and Rectum Adenocarcinoma, respectively) [12]. GEO datasets were downloaded under accession numbers GSE21510 [13] and GSE24514 [14]. The description of the selected datasets, including sample size, is presented in [S1 Table](#). All data were downloaded in June 2018. Raw expression data from GSE21510 (GPL570) and GSE24514 (GPL96) studies were normalised using the RMA method implemented in the *affy Bioconductor* R package [15] see [S1A and S1B Fig](#). Harmonized data (hg38 alignment) of RNA-Seq raw counts from the selected TCGA studies (IlluminaHiSeq) were downloaded and preprocessed using the *TCGA-biolinks Bioconductor* R package [16]. The GC-content method was used to normalise the data and quantile filtering was adopted to avoid biased gene expression values [16] (see [S1C and S1D Fig](#)). After the normalisation and filtering processes, we log-transformed the data to compare *TULP3* gene expression between tumoral and adjacent non-tumoral tissues (NT). Principal component analysis was performed in each study to filter possibly biased samples (see [S2 Fig](#)).

Statistical analyses

Data normality assumptions were verified in all studies and the statistical tests chosen to compare *TULP3* levels were: two-tailed Mann-Whitney-Wilcoxon test (for GSE21510, COAD, and READ studies), and two-tailed *t* test (for GSE24514 study).

Survival analysis was performed in COAD and READ studies using *survival Bioconductor* R package [17]. To compare the prognostic value of *TULP3* gene expression, we used the median to dichotomize the cohort. Kaplan-Meier method was used to estimate survival curves and log-rank test was used to compare the overall survival among groups. Hazard Ratios (HR) with 95% confidence intervals were calculated and *P values* from likelihood ratio tests were considered.

To measure the association between the dichotomized *TULP3* gene expression and clinicopathological information from COAD and READ studies, we performed Pearson's Chi-squared test (with Yates' continuity correction), or Fisher exact test. All data preprocessing and analyses were performed in R 3.4.2 statistical software.

Results

Bioinformatics analysis of *TULP3* gene expression

We comparatively analysed *TULP3* gene expression in GSE21510, GSE24514, COAD, and READ studies. The *TULP3* gene expression comparison between tumoral and NT samples showed significant statistical difference in all studies, with $P = 5.00e-06$, $P = 6.43e-03$, $P = 2.20e-16$ and $P = 2.13e-04$, respectively (Fig 1A–1D). In the TCGA COAD and READ studies, we observed higher fold-changes of *TULP3* gene expression in tumour samples relative to NT samples ($\log_{2}FC = 0.804$ and $\log_{2}FC = 0.590$, respectively). Additionally, *TULP3* expression profile from GSE21501 study was not able to differentiate primary CRC from metastatic (see S3 Fig).

Survival analysis

In order to verify the prognostic value of *TULP3* gene expression, we performed survival analysis using the TCGA COAD and READ datasets, since this information was readily available. We used the median to dichotomize groups classified as high and low gene expression (normalised and quantile-filtered data without log-transformation). Using this approach, no survival difference was found in patients diagnosed with COAD and READ (see S4 Fig). In order to verify whether pathologic staging (pTNM) correlates to survival, we performed Kaplan-Meier curves according to pTNM classification in COAD and READ studies. We observed that there is a difference between the early stages (I and II) versus advanced stages (III and IV). Patients diagnosed in the early stages of the disease presented better prognosis when compared to more advanced stages (see S5 Fig). The survival analysis of staging (I, II, III and IV) of COAD and READ is depicted in S6 Fig. In READ study, the Cox proportional hazards regression analysis did not converge, therefore there is no estimates of P values, neither HR. Additionally, we performed survival curves based on *TULP3* gene expression and pTNM classification, to analyse the association between these variables in COAD and READ studies. To increase the sample size, we analysed early (I and II) and advanced stages (III and IV). We used the median of *TULP3* to dichotomize the groups in low and high expression. No survival difference in *TULP3* gene expression was observed in patients with early and advanced stages (see S7 and S8 Figs).

Association of gene expression and clinical data

Chi-squared analysis was performed to verify the association among dichotomized *TULP3* gene expression and the clinical variables: gender, pathologic primary tumour extension (pT), pathologic lymph node involvement (pN), pathologic distant metastasis (pM), pathologic stage (pTNM), residual tumour, history of colon polyps, lymphatic invasion, vascular invasion, cancer status and hypermutated status. As observed in the results presented in Table 1, significant association was found in lymphatic invasion (COAD) and vascular invasion (READ), in which a higher *TULP3* expression was associated to lymphatic invasion in COAD-TCGA and to vascular invasion in READ-TCGA.

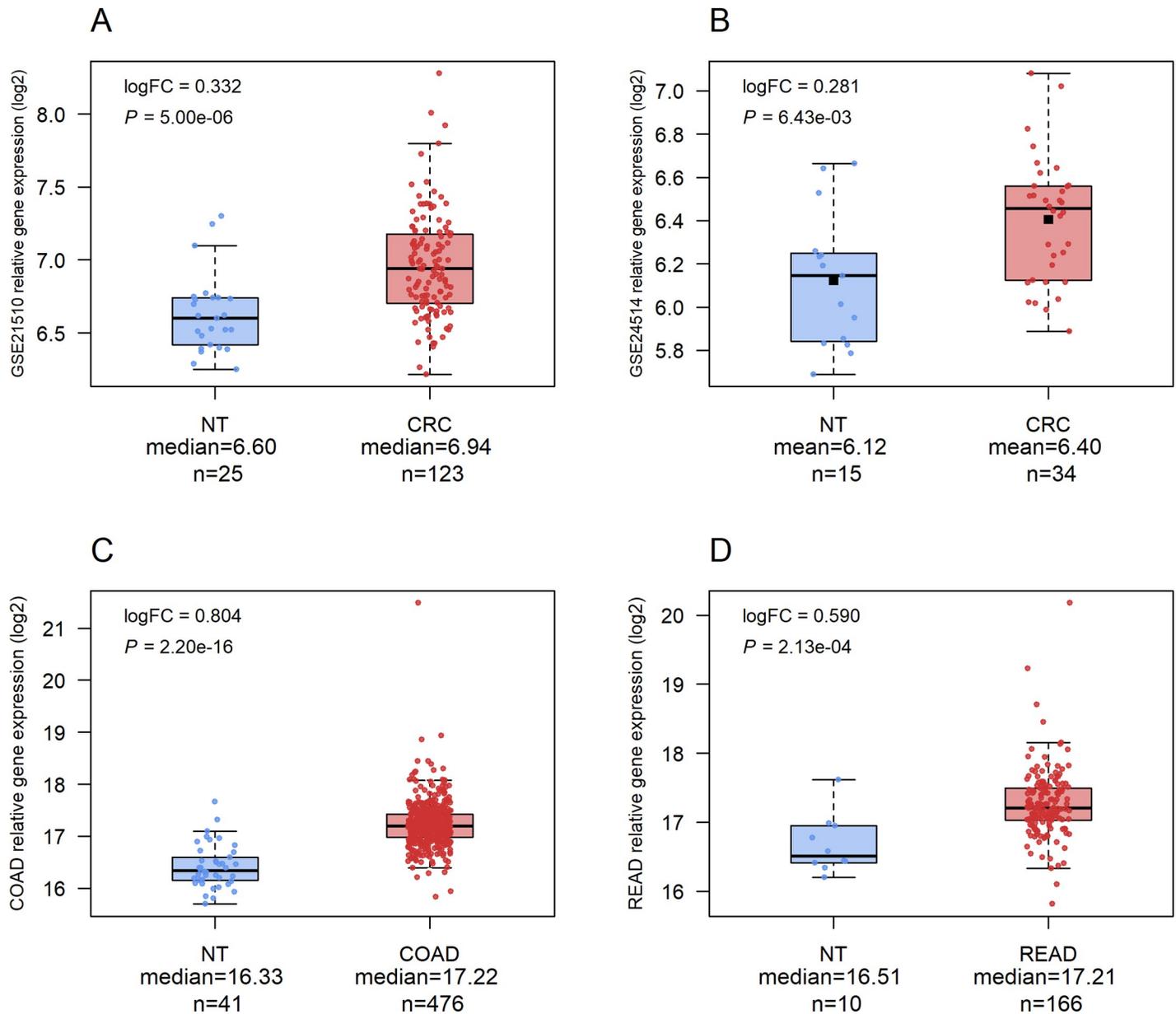


Fig 1. TULP3 gene expression comparison between groups. (A) GSE21510, (B) GSE24514, (C) COAD-TCGA, and (D) READ-TCGA. (NT) Adjacent non-tumoral tissue. (CRC) Colorectal cancer. (COAD) Colon adenocarcinoma. (READ) Rectum adenocarcinoma. (logFC) Log fold-change. Median is represented as a solid black line, whereas mean is represented as a black square.

<https://doi.org/10.1371/journal.pone.0210762.g001>

Discussion

TULP3 belongs to a family of proteins that possess a *tubby* domain in C-terminal, which exhibits a conserved nuclear localization and allows it to act as a transcription factor [18]. In a study of mammalian development [19], the authors have demonstrated the important role of *Tulp3* gene since its expression is ubiquitous throughout embryonic mice development and knockout of *Tulp3* leads to neural tube defects and embryonic lethality.

The Tulp3 protein negatively regulates the Hedgehog pathway in the absence of Sonic Hedgehog (Shh) ligand in mouse embryo at primary cilium (a microtubule-based membrane

Table 1. Association of dichotomized TULP3 gene expression with clinical variables of individuals diagnosed with COAD and READ.

| | COAD dichotomized TULP3 expression | | | READ dichotomized TULP3 expression | | |
|--------------------------------|------------------------------------|-------------------|---------------|------------------------------------|-------------------|---------------|
| | High exp. (%) | Low exp. (%) | P value | High exp. (%) | Low exp. (%) | P value |
| Gender | | | | | | |
| Female | 119 (53.60) | 103 (46.40) | 0.142† | 33 (44.00) | 42 (56.00) | 0.238† |
| Male | 117 (46.43) | 135 (53.57) | | 49 (54.44) | 41 (45.56) | |
| pT | | | | | | |
| T1 or T2 | 25 (51.02) | 24 (48.98) | 1.000† | 12 (60.00) | 8 (40.00) | 1.000† |
| T3 or T4 | 72 (51.80) | 67 (48.20) | | 29 (59.18) | 20 (40.82) | |
| pN | | | | | | |
| N0 | 68 (53.97) | 58 (46.03) | 1.000† | 27 (60.00) | 18 (40.00) | 0.676† |
| N+ | 43 (53.75) | 37 (46.25) | | 15 (53.27) | 13 (46.73) | |
| pM | | | | | | |
| M0 | 97 (56.72) | 74 (43.28) | 0.317† | 36 (58.06) | 26 (41.94) | 1.000* |
| M1 | 15 (45.45) | 18 (54.55) | | 6 (54.54) | 5 (45.46) | |
| pTNM | | | | | | |
| Stage I or II | 65 (54.17) | 55 (45.83) | 1.000† | 25 (58.14) | 18 (41.86) | 0.995† |
| Stage III or IV | 46 (54.12) | 39 (45.88) | | 16 (55.17) | 13 (44.83) | |
| Residual tumour | | | | | | |
| R0 | 94 (53.71) | 81 (46.29) | 0.614† | 35 (57.38) | 26 (42.62) | 1.000* |
| R1 or R2 | 10 (45.45) | 12 (54.55) | | 6 (54.54) | 5 (45.46) | |
| History of colon polyps | | | | | | |
| NO | 55 (57.89) | 40 (42.11) | 0.371† | 33 (61.11) | 21 (38.89) | 0.440† |
| YES | 51 (50.50) | 50 (49.50) | | 9 (47.37) | 10 (52.63) | |
| Lymphatic invasion | | | | | | |
| NO | 40 (44.94) | 49 (55.06) | 0.024† | 16 (53.33) | 14 (46.67) | 0.661† |
| YES | 63 (62.38) | 38 (37.62) | | 24 (61.54) | 15 (38.46) | |
| Vascular invasion | | | | | | |
| NO | 63 (49.22) | 65 (50.78) | 0.195† | 21 (45.65) | 25 (54.35) | 0.029† |
| YES | 29 (61.70) | 18 (38.30) | | 15 (78.95) | 4 (21.05) | |
| Cancer status | | | | | | |
| Tumor free | 19 (48.72) | 20 (51.28) | 0.589† | 1 (50.00) | 1 (50.00) | 1.000* |
| With tumor | 92 (55.09) | 75 (44.91) | | 41 (57.75) | 30 (42.25) | |
| Hypermutated status | | | | | | |
| NO | 66 (52.80) | 59 (72.20) | 0.640† | 37 (56.92) | 28 (43.08) | 1.000* |
| YES | 19 (59.37) | 13 (40.63) | | 2 (66.67) | 1 (33.33) | |
| Age mean (min-max) | 67.6 (31.2–90.0) | | | 66.2 (31.2–90.0) | | |

(pT) Pathologic primary tumour extension. (pN) Pathologic lymph nodes involvement. (pM) Pathologic distant metastasis. (pTNM) Pathologic stage. Significant data are emphasized in bold.

† Pearson Chi-squared test.

* Fisher exact test.

<https://doi.org/10.1371/journal.pone.0210762.t001>

extension that acts as sensory signaling compartment) [20, 21]. Tulp3 is also implicated in vesicle trafficking [20, 22], since the *tubby* domain binds to PIP₂ (phosphatidylinositol 4,5-bisphosphate), a component of cell membranes [23]. Mukhopadhyay *et al.* (2010) [22] highlighted the role of Tulp3 and intraflagellar transport complex A (IFT-A) in trafficking of a subset of G protein-coupled receptors (GPCRs) to cilia, in mouse embryo. Also, Santagata *et al.* (2001)

[23] suggested that Tulp3 may play a role in GPCR signaling and gene expression regulation in Neuro-2A cell.

Several studies describe the important role of Tulp3 in embryo development however, few studies have been carried out investigating its possible role in carcinogenesis. Recently, Qian *et al.* [24] used radiogenomics analysis to identify genes associated with risk prediction in glioblastoma multiforme and *TULP3* was involved. In another study, using an *in silico* approach, increased *TULP3* gene expression was associated with poor prognosis in PDAC patients, but these results were not yet validated in clinical samples [5].

We explored the profile of *TULP3* gene expression in four independent datasets related to colorectal tissues. In all of them, we observed a statistically significant increase in *TULP3* expression levels in colorectal samples compared to NT samples, despite the small number compared to tumour samples. To classify the NT specimens, some criteria must be followed, for example, NT tissue collection should be more than 2cm away from the resected lesion margin and should not contain tumour cells in histopathological analysis [25] and, sometimes, no adjacent non-tumoral tissue is available for collection. However, even with this issue is worthy to study differences in NT and tumours samples. Recent works that used TCGA and GEO publicly available datasets also found relevant results despite the difference in sample sizes [26–28]. Taking together, these results indicate a possible role of *TULP3* in the diagnosis of CRC, however, additional studies are needed to confirm these preliminary findings. Another point to consider is that using only tumour and adjacent non-tumoral tissue we possibly added bias caused by not analysing true normal samples. Although histologically NT samples are usually considered to be normal, they could present gene expression alterations [29].

Availability of survival and clinicopathological information in the TCGA datasets allowed us to investigate the prognostic impact of *TULP3* gene expression in COAD and READ patients. Although no association of *TULP3* expression profiles and survival was identified, higher levels of *TULP3* were associated with vascular and lymphatic invasion in our study. The histological identification of lymphatic and vascular invasion has long been recognized as a potential prognostic indicator and predictor of patient outcome [30, 31]. Some studies reported that lymphatic invasion is associated with an increased risk of regional lymph node metastases [32–34] and vascular invasion with risk of distant metastases [35].

The clinicopathological characteristics, especially the tumour-node-metastasis (TNM) staging and lymphovascular invasion, are related to CRC prognostic factors. Recently, molecular markers are gaining more attention as predictive of a worse outcome, an example are the members of Wnt signaling pathway, which are known to play an important role in CRC progression [36, 37]. A molecular marker associated to tumour invasion is the β -catenin, which plays key roles in the Wnt/ β -catenin signaling pathway, as well as in cellular junctions with E-cadherin. The disruption of adherens junctions is implicated in epithelial-to-mesenchymal transition (EMT) alterations, which in turn may promote tumour spreading and, consequently, metastasis. The Wnt/ β -catenin signaling pathway is also related to initiation and progression of CRC [37, 38]. A recent study, identified Tulp3 as a direct target of canonical β -catenin signaling in cerebral cortex of mice [39]. Taken together, these observations led us to hypothesize that the association found in our study between *TULP3* gene expression and, lymphatic and vascular invasion may be due to an increased activity of the β -catenin pathway. Another point to consider is that CRC is a heterogeneous and complex disease, and many studies classify both colon and rectum tumours as a common entity [40]. Evidences suggest that grouping these anatomically distinct diseases could be a clinical and biological oversimplification, therefore it would be important to study colon and rectum cancers as separated diseases, which includes differentiating them at the molecular level [40].

Conclusions

To the best of our knowledge, this preliminary study is the first one to describe increased *TULP3* gene expression in CRC compared to adjacent non-tumoral tissue, and to associate increased gene expression to vascular and lymphatic invasion in CRC. These results point to a possible role of *TULP3* as a diagnostic and prognostic biomarker in colorectal cancer.

Supporting information

S1 Table. Description of the datasets selected to analyze *TULP3* gene expression, including the classification and sample size, and the technique employed to quantify the transcripts.

(NT) Adjacent non-tumoral tissue. (CRC) Colorectal cancer. (COAD) Colon adenocarcinoma. (READ) Rectum adenocarcinoma. (DOCX)

S1 Fig. Normalisation plots. (A) GSE21510. (B) GSE24514. (C) COAD-TCGA. (D) READ-TCGA. Blue boxplots correspond to adjacent non-tumoral samples (NT) and the red ones correspond to colorectal cancer (CRC) samples.

(TIFF)

S2 Fig. Principal component analysis (PCA) plots. (A) GSE21510. (B) GSE24514. (C) COAD-TCGA. (D) READ-TCGA. (PC) Principal component. (NT) Adjacent non-tumoral tissue. (CRC) Colorectal cancer. (COAD) Colon adenocarcinoma. (READ) Rectum adenocarcinoma.

(TIFF)

S3 Fig. *TULP3* gene expression comparison between groups of GSE21501. *TULP3* expression profile from GSE21501 study. (NT) Adjacent non-tumoral tissue. (CRC) Colorectal cancer. Median is represented as a solid line. Equal letters above the boxplots indicate no statistical difference among the groups. We performed Kruskal-Wallis test followed by Benjamini-Hochberg correction for multiple comparisons.

(TIFF)

S4 Fig. Survival probabilities in TCGA studies. We used the median to dichotomize the groups classified as high and low gene expression (normalised and quantile-filtered data without log-transformation). (A) Comparison of *TULP3* gene expression in Colon adenocarcinoma (COAD). (B) Comparison of *TULP3* gene expression in Rectum adenocarcinoma (READ). The x-axis corresponds to overall survival in months. (HR) Hazard ratio.

(TIFF)

S5 Fig. Survival probabilities of early and advanced stages in TCGA studies. We grouped patients classified in stages I and II as early stage and, patients classified in stages III and IV as advanced stage. (A) Comparison of survival between early *versus* advanced stages in Colon adenocarcinoma (COAD). (B) Comparison of survival between early *versus* advanced stages in Rectum adenocarcinoma (READ). The x-axis corresponds to overall survival in months. (HR) Hazard ratio.

(RAR)

S6 Fig. Survival probabilities of pathologic staging (pTNM) in TCGA studies. (A) Comparison of survival among pTNM stages in Colon adenocarcinoma (COAD). (B) Comparison of survival among pTNM stages in Rectum adenocarcinoma (READ). The x-axis corresponds to overall survival in months. (HR) Hazard ratio. In READ study the Cox proportional hazards

regression analysis did not converged, therefore there is no estimates of P values neither HR. (RAR)

S7 Fig. Survival probabilities of *TULP3* gene expression in early and advanced stages in COAD-TCGA study. We used the median of *TULP3* to dichotomize the groups in low and high expression for early and advanced stages. (A) Comparison of survival between early stages (I and II). (B) Comparison of survival between advanced stages (III and IV). The x-axis corresponds to overall survival in months. (HR) Hazard ratio. (RAR)

S8 Fig. Survival probabilities of *TULP3* gene expression in early and advanced stages in READ-TCGA study. We used the median of *TULP3* to dichotomize the groups in low and high expression for early and advanced stages. (A) Comparison of survival between early stages (I and II). (B) Comparison of survival between advanced stages (III and IV). The x-axis corresponds to overall survival in months. (HR) Hazard ratio. (RAR)

Author Contributions

Conceptualization: Ivaine Taís Sauthier Sartor, Mariana Recamonde-Mendoza, Patricia Ashton-Prolla.

Formal analysis: Ivaine Taís Sauthier Sartor, Mariana Recamonde-Mendoza.

Funding acquisition: Mariana Recamonde-Mendoza, Patricia Ashton-Prolla.

Supervision: Mariana Recamonde-Mendoza, Patricia Ashton-Prolla.

Writing – original draft: Ivaine Taís Sauthier Sartor.

Writing – review & editing: Mariana Recamonde-Mendoza, Patricia Ashton-Prolla.

References

1. GLOBOCAN 2012 (IARC). Estimated cancer incidence, mortality, and prevalence worldwide in 2012. From: <http://globocan.iarc.fr/old/FactSheets/cancers/stomach-new.asp> Accessed 26 August 2018.
2. Costa NR, Costa RMG, Medeiros R. A viral map of gastrointestinal cancers. *Life Sciences*. 2018; 199: 188–200. <https://doi.org/10.1016/j.lfs.2018.02.025> PMID: 29476768
3. Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med*. 2015; 21(5): 449–456. <https://doi.org/10.1038/nm.3850> PMID: 25894828
4. Gray JW, Collins C. Genome changes and gene expression in human solid tumors. *Carcinogenesis*. 2000; 21(3): 443–452. PMID: 10688864
5. Sartor ITS, Zeidán-Chuliá F, Albanus RD, Dalmolin RJS, Moreira JCF. Computational analyses reveal a prognostic impact of TULP3 as a transcriptional master regulator in pancreatic ductal adenocarcinoma. *Mol Biosyst*. 2014; 10(6): 1461–1468. <https://doi.org/10.1039/c3mb70590k> PMID: 24668219
6. Tam PPL, Kanai-Azumay M, Kanaiz Y. Early endoderm development in vertebrates: lineage differentiation and morphogenetic function. *Curr Opin Genetics Dev*. 2003; 13(4): 393–400. [https://doi.org/10.1016/S0959-437X\(03\)00085-6](https://doi.org/10.1016/S0959-437X(03)00085-6)
7. Pinchuk IV, Mifflin RC, Saada JI, Powell DW. Intestinal Mesenchymal Cells. *Curr Gastroenterol Rep*. 2010; 12(5): 310–318. <https://doi.org/10.1007/s11894-010-0135-y> PMID: 20690004
8. Fausett SR, Klingensmith J. Compartmentalization of the foregut tube: developmental origins of the trachea and esophagus. *Dev Biol*. 2012; 1(2): 184–202. <https://doi.org/10.1002/wdev.12>
9. Jennings RE, Berry AA, Strutt JP, Gerrard DT, Hanley NA. Human pancreas development. *Devel*. 2015; 142(18): 3126–3137. <https://doi.org/10.1242/dev.120063>
10. Kim TH, Shivdasani RA. Stomach development, stem cells and disease. *Devel*. 2016; 143(4): 554–565. <https://doi.org/10.1242/dev.124891>

11. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002; 30(1): 207–210. PMID: [11752295](#)
12. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature.* 2014; 513(7517): 202–209. <https://doi.org/10.1038/nature13480> PMID: [25079317](#)
13. Tsukamoto S, Ishikawa T, Iida S, Ishiguro M, Mogushi K, Mizushima H, et al. Clinical significance of osteoprotegerin expression in human colorectal cancer. *Clin Cancer Res.* 2011; 17(8): 2444–2450. <https://doi.org/10.1158/1078-0432.CCR-10-2884> PMID: [21270110](#)
14. Alhopuro P, Sammalkorpi H, Niittymäki I, Biström M, Raitila A, Saharinen J, et al. Candidate driver genes in microsatellite-unstable colorectal cancer. *Int J Cancer.* 2012; 130(7): 1558–1566. <https://doi.org/10.1002/ijc.26167> PMID: [21544814](#)
15. Gautier L, Cope L, Bolstad BM, Irizarry, R.A. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics.* 2004; 20(3): 307–315. <https://doi.org/10.1093/bioinformatics/btg405> PMID: [14960456](#)
16. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Carolini D, et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016; 44(8): e71. <https://doi.org/10.1093/nar/gkv1507> PMID: [26704973](#)
17. Therneau TM, Grambsch PM. Estimating the survival and hazard functions. In: Dietz K, Gail M, Krickeberg K, Samet J, Tsiatis A, editors. *Modeling Survival Data: Extending the Cox Model.* New York Springer; 2000. p. 7–31. <https://doi.org/10.1007/978-1-4757-3294-8>
18. Boggon TJ, Shan WS, Santagata S, Myers SC, Shapiro L. Implication of tubby proteins as transcription factors by structure-based functional analysis. *Science.* 1999; 286(5447): 2119–2125. <http://doi.org/10.1126/science.286.5447.2119> PMID: [10591637](#)
19. Ikeda A, Ikeda S, Gridley T, Nishina PM, Naggert JK. Neural tube defects and neuroepithelial cell death in Tulp3 knockout mice. *Hum Mol Genet.* 2011; 10(12): 1325–1334. <https://doi.org/10.1093/hmg/10.12.1325>
20. Norman RX, Ko HW, Huang V, Eun CM, Abler LL, Zhang Z, et al. Tubby-like protein 3 (TULP3) regulates patterning in the mouse embryo through inhibition of Hedgehog signaling. *Hum Mol Genet.* 2009; 18(10): 1740–1754. <https://doi.org/10.1093/hmg/ddp113> PMID: [19286674](#)
21. Cameron DA, Pennimpede T, Petkovich M. Tulp3 is a critical repressor of mouse hedgehog signaling. *Dev Dyn.* 2009; 238(5): 1140–1149. <https://doi.org/10.1002/dvdy.21926> PMID: [19334287](#)
22. Mukhopadhyay S, Wen X, Chih B, Nelson CD, Lane WS, Scales SJ, et al. TULP3 bridges the IFT-A complex and membrane phosphoinositides to promote trafficking of G protein-coupled receptors into primary cilia. *Genes Dev.* 2010; 24(19): 2180–2193. <https://doi.org/10.1101/gad.1966210> PMID: [20889716](#)
23. Santagata S, Boggon TJ, Baird CL, Gomez CA, Zhao J, Shan WS, et al. G-protein signaling through tubby proteins. *Science.* 2001; 292(5524): 2041–2050. <https://doi.org/10.1126/science.1061233> PMID: [11375483](#)
24. Qian X, Tan H, Zhang J, Liu K, Yang T, Wang M, et al. Identification of biomarkers for pseudo and true progression of GBM based on radiogenomics study. *Oncotarget.* 2016; 7(34): 55377–55394. <https://doi.org/10.18632/oncotarget.10553> PMID: [27421136](#)
25. Aran D, Camarda R, Odegaard J, Paik H, Oskotsky B, Krings G, Goga A, et al. Comprehensive analysis of normal adjacent to tumor transcriptomes. *Nat Commun.* 2017; 8(1): 1077. <https://www.nature.com/articles/s41467-017-01027-z> PMID: [29057876](#)
26. Xing Y, Zhao Z, Zhu Y, Zhao L, Zhu A, Piao D. Comprehensive analysis of differential expression profiles of mRNAs and lncRNAs and identification of a 14-lncRNA prognostic signature for patients with colon adenocarcinoma. *Oncol Rep.* 2018; 39(5): 2365–2375. <https://doi.org/10.3892/or.2018.6324> PMID: [29565464](#)
27. Han T, Lin J, Wang Y, Fan Q, Sun H, Tao Y, et al. Forkhead box D1 promotes proliferation and suppresses apoptosis via regulating polo-like kinase 2 in colorectal cancer. *Biomed Pharmacother.* 2018; 103: 1369–1375. <https://doi.org/10.1016/j.biopha.2018.04.190> PMID: [29864920](#)
28. Li C, Shen Z, Zhou Y, Yu W. Independent prognostic genes and mechanism investigation for colon cancer. *Biol Res.* 2018; 51(1): 10. <https://doi.org/10.1186/s40659-018-0158-7> PMID: [29653552](#)
29. Assumpção PP, Santos SEB, Santos AKR, Demachki S, Khayat AS, Ishak G, et al. The adjacent tumor sample trap. *Gastric Cancer.* 2016; 19(3): 1024–1025. <http://dx.doi.org/10.1007/s10120-015-0539-3> PMID: [26359158](#)
30. Fujii T, Sutoh T, Morita H, Yajima R, Yamaguchi S, Tsutsumi S, et al. Vascular invasion, but not lymphatic invasion, of the primary tumor is a strong prognostic factor in patients with colorectal cancer. *Anti-cancer Res.* 2014; 34(6): 3147–3152. PMID: [24922686](#)

31. Harris EI, Lewin DN, Wang HL, Lauwers GY, Srivastava A, Shyr Y, et al. Lymphovascular invasion in colorectal cancer: an interobserver variability study. *Am J Surg Pathol*. 2008; 32(12): 1816–1821. <https://dx.doi.org/10.1097%2FAPAS.0b013e3181816083> PMID: 18779725
32. Coverlizza S, Risio M, Ferrari A, Fenoglio-Preiser CM, Rossini FP. Colorectal adenomas containing invasive carcinoma: pathologic assessment of lymph node metastatic potential. *Cancer*. 1989; 64(9): 1937–1947. PMID: 2477139
33. Muller S, Chesner IM, Egan MJ, Rowlands DC, Collard MJ, Swarbrick ET, et al. Significance of venous and lymphatic invasion in malignant polyps of the colon and rectum. *Gut*. 1989; 30(10): 1385–1391. PMID: 2583564
34. Nivatvongs S, Rojanasakul A, Reiman HM, Dozois RR, Wolff BG, Pemberton JH, et al. The risk of lymph node metastasis and colorectal polyps with invasive carcinoma. *Dis Colon Rectum*. 1991; 34(4): 323–328. PMID: 1848810
35. Ouchi K, Sugawara T, Ono H, Fujiya T, Kamiyama Y, Kakugawa Y, et al. Histologic features and clinical significance of venous invasion in colorectal carcinoma with hepatic metastasis. *Cancer*. 1996; 78(11): 2313–2317. [https://doi.org/10.1002/\(SICI\)1097-0142\(19961201\)78:11<2313::AID-CNCR7>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1097-0142(19961201)78:11<2313::AID-CNCR7>3.0.CO;2-N) PMID: 8941000
36. Koumariou A, Tzeveleki I, Mekras D, Eleftheraki AG, Bobos M, Wirtz R, et al. Prognostic markers in early-stage colorectal cancer: Significance of TYMS mRNA expression. *Anticancer Res*. 2014; 34(9): 4949–4962. PMID: 25202077
37. Yoshida N, Kinugasa T, Ohshima K, Yuge K, Ohchi T, Fujino S, et al. Analysis of Wnt and β -catenin Expression in Advanced Colorectal Cancer. *Anticancer Res*. 2015; 35(8): 4403–4410. PMID: 26168479
38. Umbreit C, Flanjak J, Weiss C, Erben P, Aderhold C, Faber A, et al. Incomplete Epithelial–Mesenchymal Transition in p-16 positive squamous cell carcinoma cells correlates with β -catenin expression. *Anticancer Res*. 2014; 34(12): 7061–7070. PMID: 25503133
39. Nakagawa N, Li J, Yabuno-Nakagawa K, Eom TY, Cowles M, Mapp T, et al. APC sets the Wnt tone necessary for cerebral cortical progenitor development. *Genes Dev*. 2017; 31(16): 1679–1692. <https://doi.org/10.1101/gad.302679.117> PMID: 28916710
40. Sanz-Pamplona R, Cordero D, Berenguer A, Lejbkovicz F, Rennert H, Salazar R, et al. Gene expression differences between colon and rectum tumors. *Clin Cancer Res*. 2011; 17(23): 7303–7312. <https://doi.org/10.1158/1078-0432.CCR-11-1570> PMID: 21976543

Supporting Information

S1 Table. Description of the datasets selected to analyze *TULP3* gene expression, including the classification and sample size, and the technique employed to quantify the transcripts. (NT) Adjacent non-tumoral tissue. (CRC) Colorectal cancer. (COAD) Colon adenocarcinoma. (READ) Rectum adenocarcinoma.

| Selected datasets | Classification of samples (n) | Technique used to quantify transcripts (platform) |
|-------------------|-------------------------------|---|
| GSE21510 | NT (25) | Microarray |
| | CRC (123) | Affymetrix Human Genome U133 Plus 2 (GPL570) |
| GSE24514 | NT (15) | Microarray |
| | CRC (34) | Affymetrix Human Genome U133 A (GPL96) |
| COAD-TCGA | NT (41) | Harmonized data |
| | COAD (476) | (hg38 alignment) |
| READ-TCGA | NT (10) | RNA-Seq |
| | READ (166) | Harmonized data (hg38 alignment) |
| | | RNA-Seq |

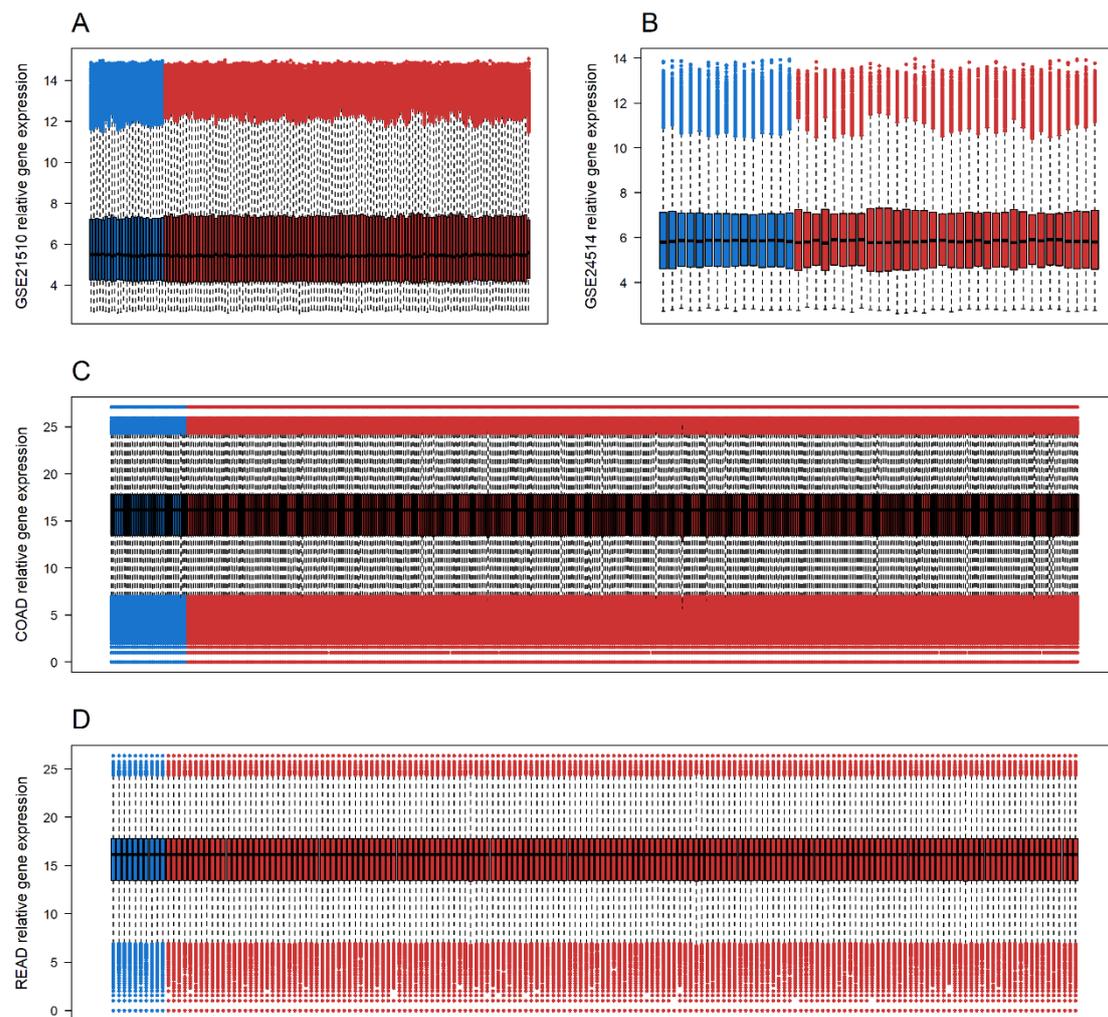


Fig. S1. Normalisation plots.

(A) GSE21510. (B) GSE24514. (C) COAD-TCGA. (D) READ-TCGA. Blue boxplots correspond to adjacent non-tumoral samples (NT) and the red ones correspond to colorectal cancer (CRC) samples.

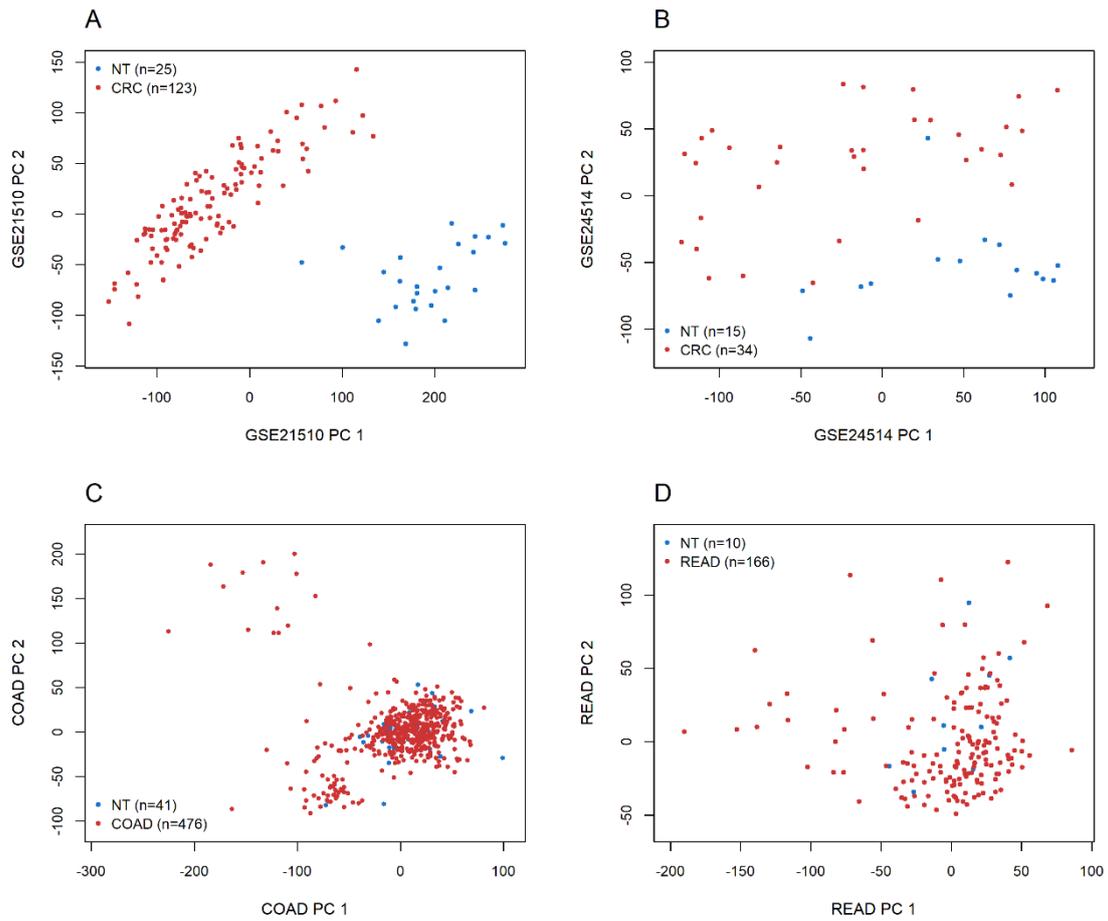


Fig. S2. Principal component analysis (PCA) plots.

(A) GSE21510. (B) GSE24514. (C) COAD-TCGA. (D) READ-TCGA. (PC) Principal component. (NT) Adjacent non-tumoral tissue. (CRC) Colorectal cancer. (COAD) Colon adenocarcinoma. (READ) Rectum adenocarcinoma.

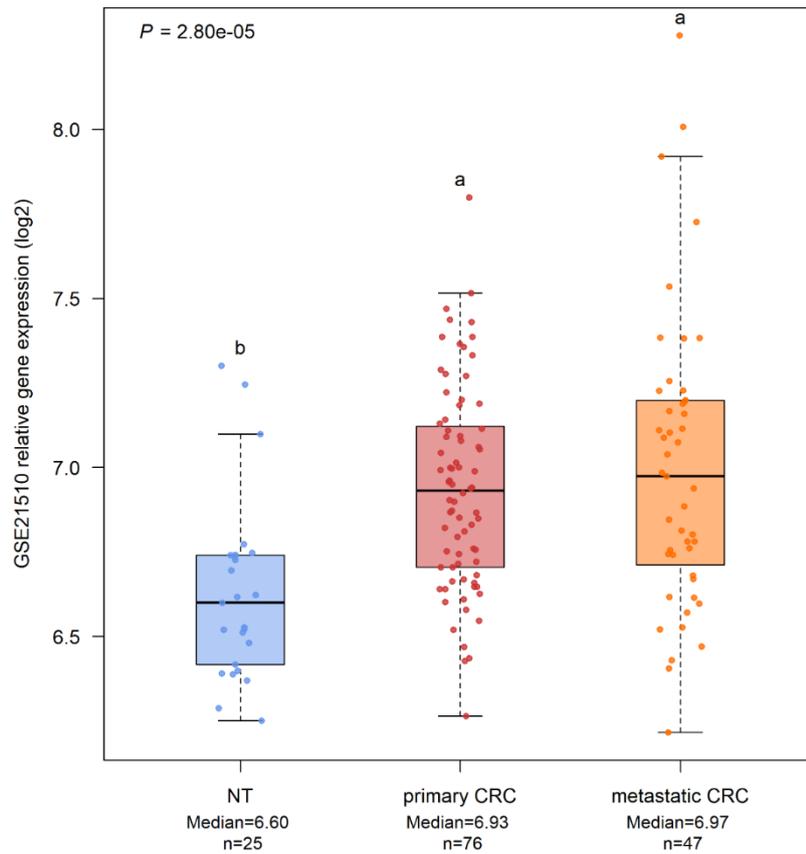


Fig. S3. *TULP3* gene expression comparison between groups of GSE21501.

TULP3 expression profile from GSE21501 study. (NT) Adjacent non-tumoral tissue. (CRC) Colorectal cancer. Median is represented as a solid line. Equal letters above the boxplots indicate no statistical difference among the groups. We performed Kruskal-Wallis test followed by Benjamini-Hochberg correction for multiple comparisons.

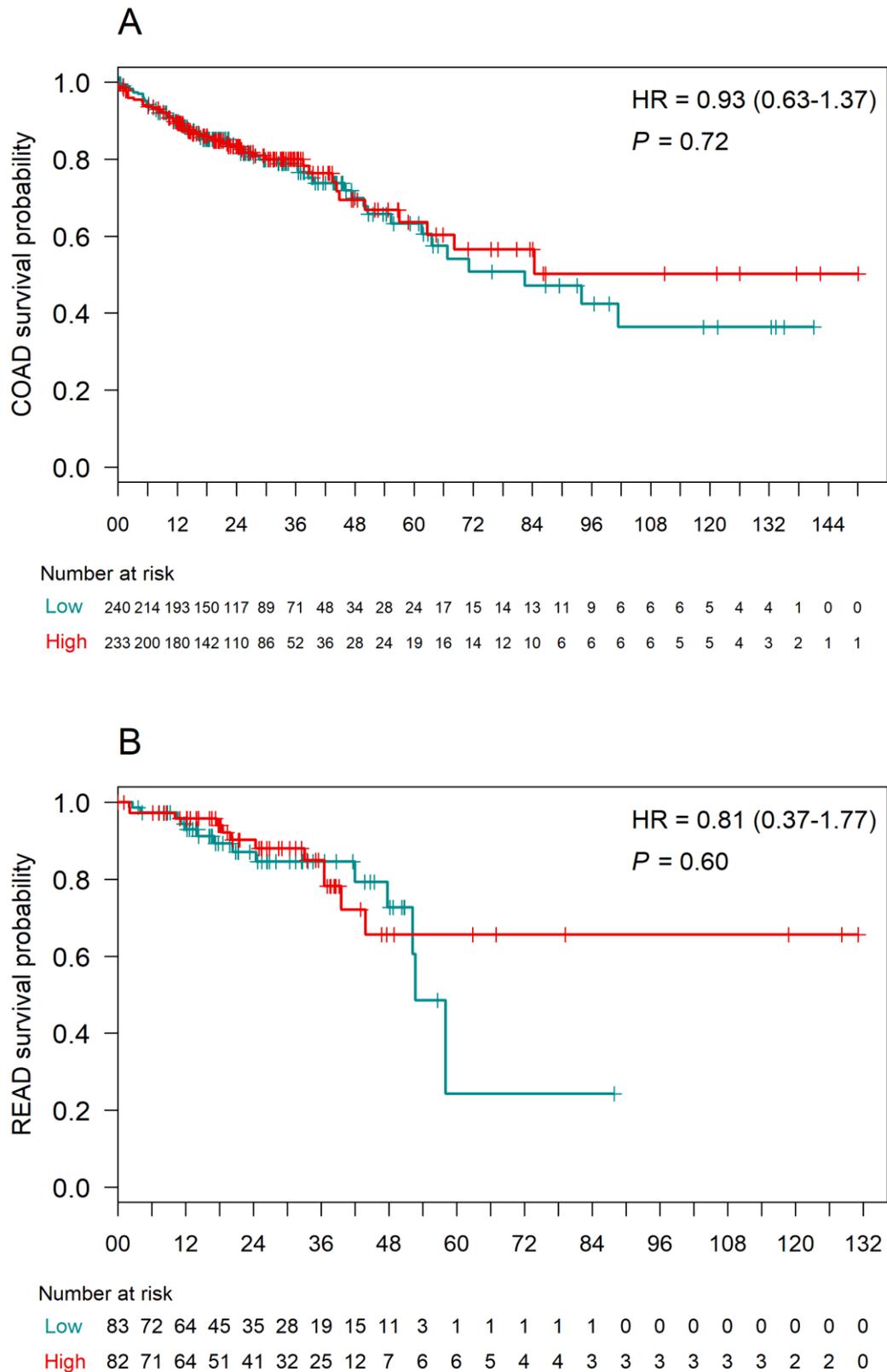


Fig. S4. Survival probabilities in TCGA studies.

We used the median to dichotomize the groups classified as high and low gene expression (normalised and quantile-filtered data without log-transformation). **(A)** Comparison of *TULP3* gene expression in Colon adenocarcinoma (COAD). **(B)** Comparison of *TULP3* gene expression in Rectum adenocarcinoma (READ). The x-axis corresponds to overall survival in months. (HR) Hazard ratio.

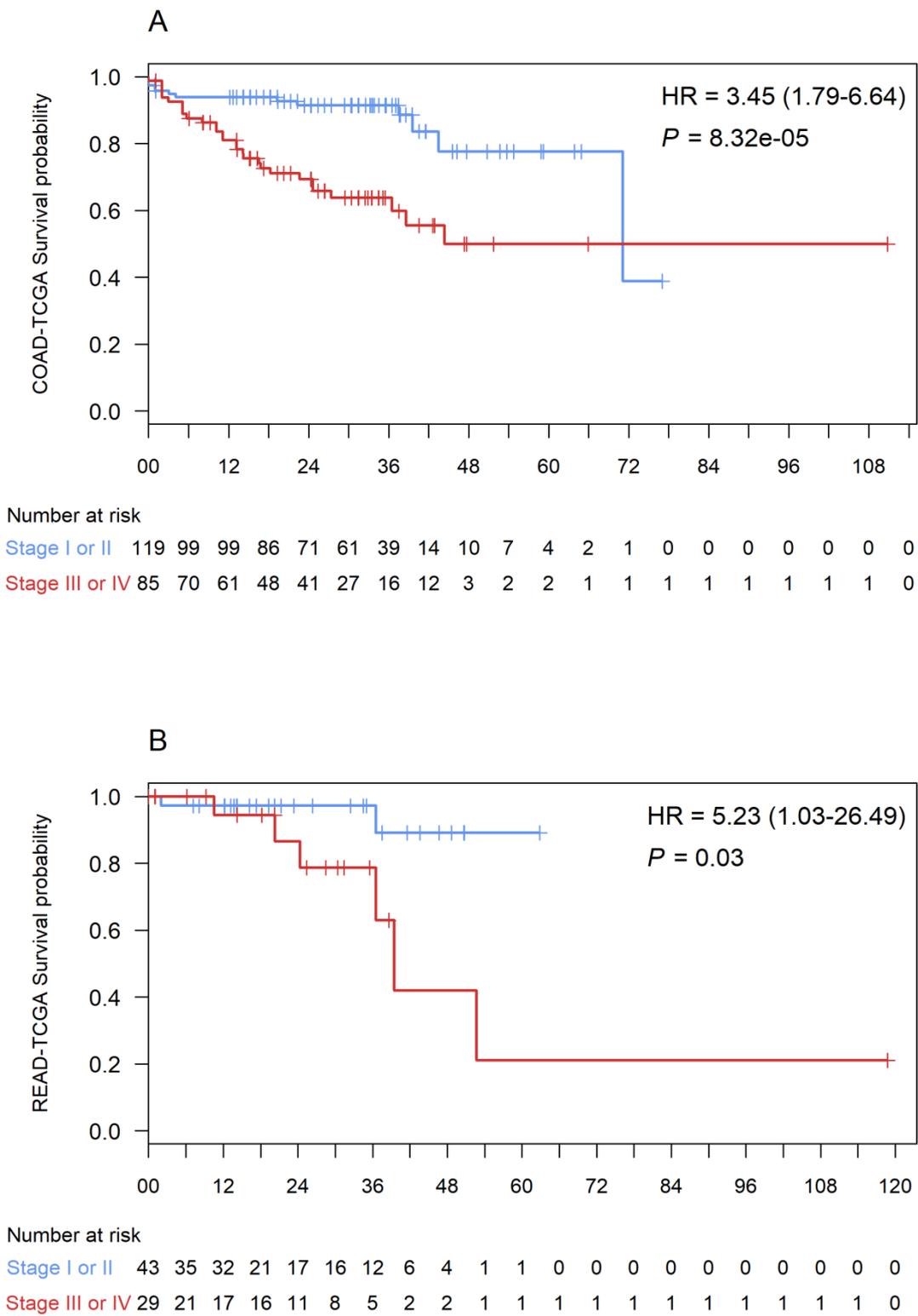


Fig. S5. Survival probabilities of early and advanced stages in TCGA studies.

We grouped patients classified in stages I and II as early stage and, patients classified in stages III and IV as advanced stage. **(A)** Comparison of survival between early *versus* advanced stages in Colon adenocarcinoma (COAD). **(B)** Comparison of survival between early *versus* advanced stages in Rectum adenocarcinoma (READ). The x-axis corresponds to overall survival in months. (HR) Hazard ratio.

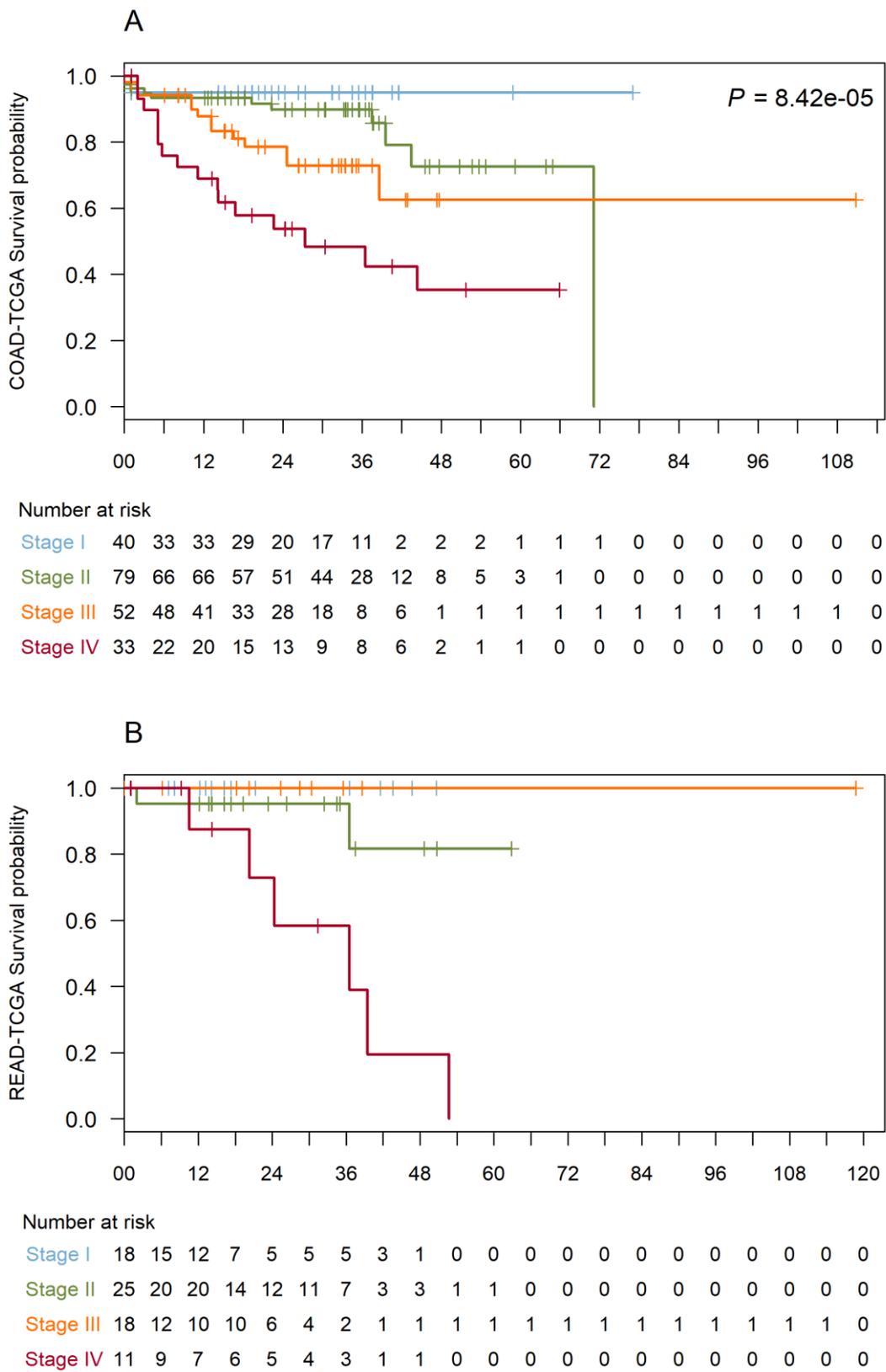


Fig. S6. Survival probabilities of pathologic staging (pTNM) in TCGA studies.

(A) Comparison of survival among pNTM stages in Colon adenocarcinoma (COAD). (B) Comparison of survival among pNTM stages in Rectum adenocarcinoma (READ). The x-axis corresponds to overall survival in months. (HR) Hazard ratio. In READ study the Cox proportional hazards regression analysis did not converged, therefore there is no estimates of *P* values neither HR.

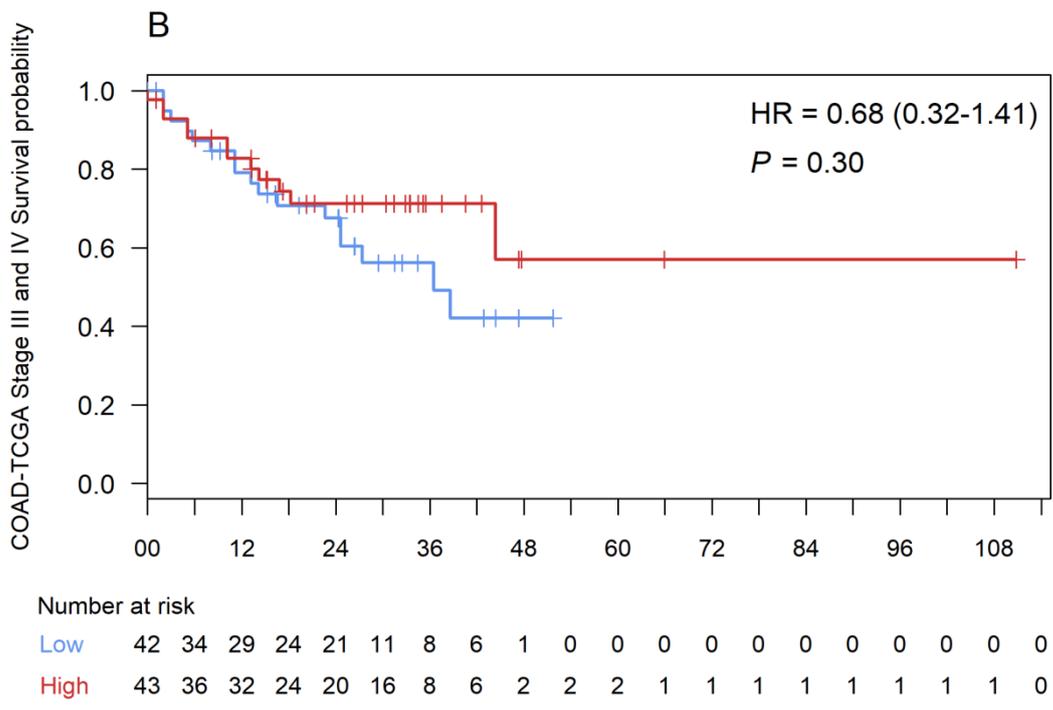
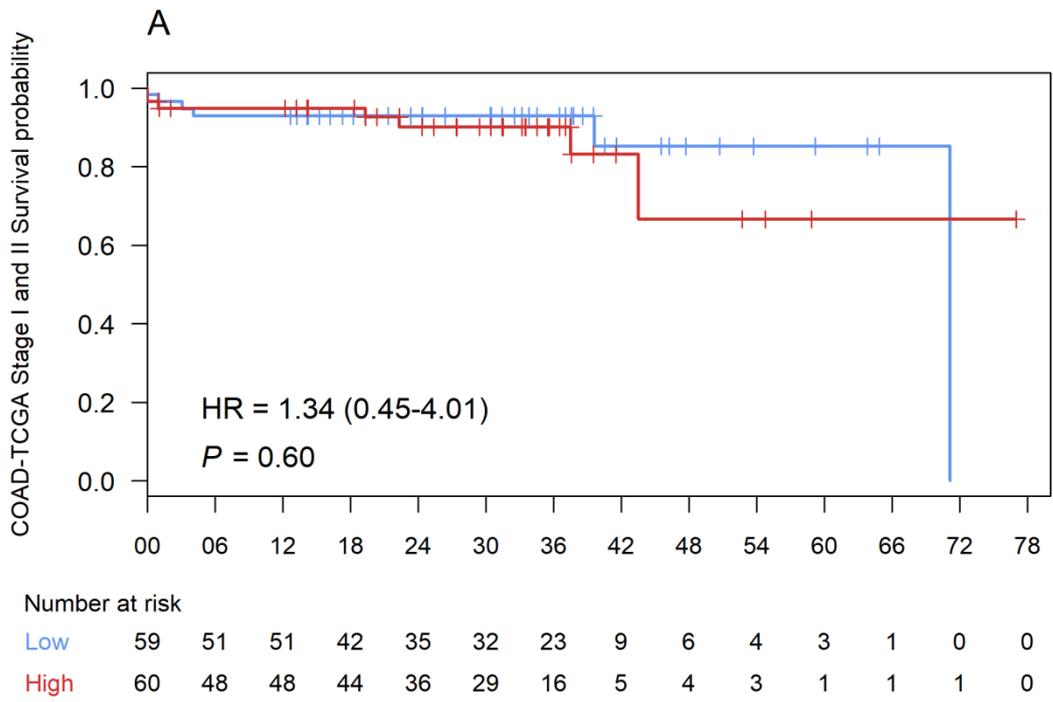


Fig. S7. Survival probabilities of *TULP3* gene expression in early and advanced stages in COAD-TCGA study.

We used the median of *TULP3* to dichotomize the groups in low and high expression for early and advanced stages. **(A)** Comparison of survival between early stages (I and II). **(B)** Comparison of survival between advanced stages (III and IV). The x-axis corresponds to overall survival in months. (HR) Hazard ratio.

Fig. S8. Survival probabilities of *TULP3* gene expression in early and advanced stages in READ-TCGA study.

We used the median of *TULP3* to dichotomize the groups in low and high expression for early and advanced stages. **(A)** Comparison of survival between early stages (I and II). **(B)** Comparison of survival between advanced stages (III and IV). The x-axis corresponds to overall survival in months. (HR) Hazard ratio.

Capítulo VI

Discussão Geral

Vários estudos descrevem o importante papel de *Tulp3* no desenvolvimento embrionário, no entanto, poucos estudos foram realizados investigando seu possível papel na carcinogênese. Recentemente, Qian *et al.* (2016) utilizaram a análise radiogenômica para identificar genes associados à predição de risco em glioblastoma multiforme e TULP3 estava envolvido. Em outro estudo, usando uma abordagem *in silico*, o aumento da expressão gênica de *TULP3* foi associado com prognóstico desfavorável em pacientes com ADP, entretanto esses resultados ainda não foram validados em amostras clínicas (Sartor *et al.*, 2014).

No manuscrito descrito no capítulo IV, foi explorado o perfil da expressão do gene *TULP3* em quatro conjunto de dados de expressão gênica independentes, relacionados aos tecidos colorretais. Em todos eles, observamos um aumento estatisticamente significativo nos níveis de expressão de *TULP3* nas amostras de CCR em comparação às amostras de tecido não-tumoral adjacente (NT), apesar do pequeno número comparado às amostras de tumores. Para classificar os espécimes NT, alguns critérios devem ser seguidos, por exemplo, a coleta de tecido NT deve estar a mais de 2cm da margem da lesão ressecada e não deve conter células tumorais na análise histopatológica (Aran *et al.*, 2017) e, em algumas vezes, não há tecido NT disponível para coleta. Trabalhos recentes que usaram conjuntos de dados publicamente disponíveis do TCGA e do GEO também encontraram resultados relevantes, apesar da diferença no tamanho amostral (Xing *et al.*, 2018; Han *et al.*, 2018; Li *et al.*, 2018). Desta maneira, os resultados apresentados indicam um possível papel de *TULP3* como biomarcador de diagnóstico de CCR, no entanto, estudos adicionais são necessários para confirmar esses achados preliminares. Outro ponto a ser considerado é que usando apenas amostras tumorais e NT, adicionamos, possivelmente, um viés já que não analisamos amostras verdadeiramente normais de tecido colorretal. Embora, histologicamente, as amostras de NT sejam geralmente consideradas normais, elas podem apresentar alterações na expressão gênica (Assumpção *et al.*, 2016).

A disponibilidade de informações clínicas e de sobrevida nos dados do TCGA permitiu investigar o impacto prognóstico da expressão do gene *TULP3* em pacientes diagnosticados com adenocarcinoma de cólon (COAD) e adenocarcinoma retal (READ). Embora não tenha sido identificada associação entre os perfis de expressão de *TULP3* e a sobrevida nesses tipos tumorais, níveis elevados de *TULP3* foram associados à invasão linfática em COAD e vascular em READ. A identificação histológica da invasão linfática e vascular tem sido reconhecida como um indicador prognóstico potencial e preditor do prognóstico do paciente (Harris et al., 2008; Fuji et al., 2014). Alguns estudos relataram que a invasão linfática está associada a um risco aumentado de metástases linfonodais regionais (Coverlizza et al., 1989; Muller et al., 1989; Nivatvongs et al., 1991) e invasão vascular com risco de metástases à distância (Ouchi et al., 1996).

As características clinicopatológicas, principalmente o estadiamento TNM e a invasão linfo-vascular, estão relacionadas aos fatores prognósticos do CCR. Recentemente, os marcadores moleculares estão ganhando mais atenção como preditivos de um resultado desfavorável, um exemplo são os membros da via de sinalização Wnt, que são conhecidos por desempenhar um papel importante na progressão do CCR (Koumarianou et al., 2014; Yoshida et al., 2015). Um marcador molecular associado à invasão tumoral é a β -catenina, que desempenha um papel importante na via de sinalização Wnt / β -catenina, bem como nas junções celulares com a E-caderina. A ruptura das junções aderentes está implicada nas alterações da transição epitélio-mesenquimal (EMT), que por sua vez podem promover a disseminação do tumor e, conseqüentemente, a metástase. A via de sinalização Wnt / β -catenina também está relacionada ao início e progressão do CCR (Umbreit et al., 2014; Yoshida et al., 2015). Um estudo recente identificou *Tulp3* como um alvo direto da sinalização canônica de β -catenina no córtex cerebral de camundongos (Nakagawa et al., 2017). Em conjunto, essas observações nos levaram a supor que a associação encontrada em nosso estudo entre a expressão do gene *TULP3* e a invasão linfática e

vascular pode ser devido ao aumento da atividade da via da β -catenina. Outro ponto a considerar é que o CCR é uma doença heterogênea e complexa, e muitos estudos classificam os tumores de cólon e reto como uma entidade comum (Sanz-Pamplona *et al.*, 2017). Evidências sugerem que agrupar essas doenças anatomicamente distintas poderia ser uma simplificação clínica e biológica, portanto, seria importante estudar o câncer de cólon e reto como doenças separadas, o que inclui sua diferenciação no nível molecular (Sanz-Pamplona *et al.*, 2017).

Para o CG, os resultados da análise de sobrevida baseada na expressão gênica nos levaram a supor que *TULP3* pode ter um papel no prognóstico no DGC. Como sugerido por (Ge *et al.*, 2018), devemos considerar a análise da correlação entre a superexpressão de proteínas e a sobrevida global para procurar possíveis alvos no DGC. Assim, realizamos imunomarcção em amostras de câncer gástrico provenientes de pacientes do norte do Brasil. No entanto, não observamos diferença na sobrevida considerando grupos com imunocoloração positiva e negativa para *TULP3*. O pequeno número de amostras com coloração positiva disponível, que incluiu apenas cinco pacientes com IGC e dois com DGC, provavelmente apresentou uma forte limitação ao poder estatístico dessa análise. Portanto, novos estudos com amostras maiores são necessários para confirmar o possível papel de *TULP3* como um biomarcador prognóstico no DGC.

Considerando o viés ligado ao tecido não-tumoral adjacente, realizamos análises adicionais considerando lesões gástricas como gastrite crônica (ChG) e metaplasia intestinal gástrica (GIM), além dos tecidos tumorais e NT. A expressão gênica de *TULP3* foi capaz de diferenciar os tecidos CG e NT em somente dois estudos (STAD e GSE33335) dos quatro selecionados. Curiosamente, notamos um aumento da expressão gênica de *TULP3* a partir das lesões gástricas ao CG. Considerando o aumento dos níveis de *TULP3* poderíamos supor que esse aumento esteja possivelmente associado à transformação celular da mucosa gástrica, e poderíamos

esperar seu aumento regular conforme o estadiamento, entretanto não observamos esse comportamento.

A lesão pré-maligna GIM pode ser dividida em dois histo-tipos: metaplasia intestinal completa (CIM) e metaplasia intestinal incompleta (IIM), sendo que o subtipo IIM apresenta maior probabilidade de progredir para o CG do que o CIM (Correa *et al.*, 2012; González *et al.*, 2016; Companioni *et al.*, 2017). A análise da expressão gênica utilizando o conjunto de dados GSE78523, que compreende NT e GIM (CIM e IIM, que evoluiu ou não para carcinoma gástrico), demonstrou que a IIM apresentou níveis mais elevados de *TULP3* quando comparado à CIM. Além dos histo-tipos GIM, o desenvolvimento e estabelecimento de GIM no estômago está intimamente relacionado à expressão dos fatores de transcrição específicos do intestino *CDX1* e *CDX2* (Kang *et al.*, 2011). No estudo apresentado no capítulo III, verificamos que *TULP3* apresentou correlação positiva com a expressão do gene *CDX1* em GIM (em ambos CIM e IIM), enquanto que nenhuma correlação entre *TULP3* e *CDX2* foi observada. Estudos anteriores relataram que a expressão de RNAm e proteína de *CDX1* e *CDX2* é estritamente confinada ao intestino, desde o duodeno até o reto. Embora a mucosa gástrica normal não expresse esses fatores transcricionais, a expressão aberrante de *CDX1* e *CDX2* é observada em GIM (Correa *et al.*, 2012; González *et al.*, 2013; Companioni *et al.*, 2017). Considerando apenas o *CDX2*, sua expressão está associada a vários carcinomas avançados, como o de estômago, cólon, pâncreas, vesícula biliar e pulmões (Seno *et al.*, 2002; Kim *et al.*, 2005; Kang *et al.*, 2005). Além disso, a positividade de *CDX2* também foi correlacionada à displasia, uma lesão pré-neoplásica (Mizoshita *et al.*, 2001; Bai *et al.*, 2002). Ainda, alguns estudos relataram uma redução em *CDX1* e *CDX2* no GC em comparação com GIM (Seno *et al.*, 2002; Kim *et al.*, 2005).

Tem sido proposto que *CDX1* e *CDX2* podem ter papéis diferentes na progressão do GIM. Alguns estudos relataram expressão prévia de *CDX2* no desenvolvimento de GIM, seguida por expressão tardia e fraca de *CDX1* (Eda *et al.*, 2002; Kang *et al.*, 2011).

Embora existam estudos destacando a importância do CDX2 no GIM, não há consenso dos papéis de CDX1 e CDX2. Assim, mais estudos são necessários para elucidar essa questão (Shin *et al.*, 2016). A partir desses resultados, podemos supor que *TULP3* e *CDX1* apresentam um papel relevante no desenvolvimento da lesão pré-maligna do GIM ao invés da transformação do CG.

Ainda, a fim de verificar o envolvimento da expressão da proteína TULP3 nos tecidos gástricos, realizamos imunohistoquímica em quatro blocos de TMA, compostos por amostras de ChG, GIM e CG. Com relação aos padrões de imunomarcação, observamos uma coloração citoplasmática difusa e *dot-like* na maioria das amostras analisadas (GIM e CG). Apesar de seu papel como fator de transcrição, a proteína TULP3 também está envolvida no tráfego de vesículas (Norman *et al.*, 2009; Mukhopadhyay *et al.*, 2010), uma vez que se liga a PIP₂ (Santagata *et al.*, 2001), possivelmente explicando o padrão de coloração citoplasmático observado. Por fim, encontramos uma associação significativa entre a imunocoloração positiva com as lesões GIM, entretanto não observados esse padrão com amostras de ChG e CG. Portanto, esses resultados reforçam o papel da TULP3 no GIM e, possivelmente, seu papel como biomarcador dessa lesão.

Por fim, a avaliação do perfil de expressão proteica de TULP3 em tecidos pancreáticos foi realizada em amostras de adenocarcinoma ductal pancreático (ADP), tecido pancreático não-tumoral adjacente (PNT), pancreatite crônica, lesões intraepiteliais pancreáticas (PanIN), carcinoma de ampola de Vater (CAV) e tecido periampular não-tumoral adjacente (ANT). Considerando que a abundância de RNAm não prediz com confiabilidade o conteúdo proteico (Cristescu *et al.*, 2015), este estudo teve como objetivo avaliar a expressão da proteína TULP3 em amostras de tecido pancreático por meio de imunocoloração, associando sua expressão ao desfecho e informações clínico-patológicas. Observamos uma associação entre o escore de imunorreatividade (IRS) de TULP3 positivo com PNT e PanIN e uma associação

negativa com ADP. Também avaliamos a associação com IRS de TULP3 com informações clínico-patológicas, no entanto, nenhuma associação foi verificada no estadiamento ADP, pT, pN e pM. Devido ao pequeno tamanho amostral de AVC, ANT e CP, não foi possível avaliar o perfil de imunocoloração de TULP3; portanto, são necessários mais estudos com tamanhos maiores. Esses achados indicam que TULP3 pode não estar associado à carcinogênese pancreática, mas no desenvolvimento celular de tecido pancreático não-tumoral adjacente e das lesões intraepiteliais pancreáticas

Esse trabalho apresentou algumas limitações relacionadas ao desenvolvimento dos estudos citados anteriormente como: impossibilidade de realizar quantificação de RNAm nas linhagens estabelecidas de câncer gástrico, o que permitiria a realização de um estudo de correlação entre RNAm e proteína em CG. Outra dificuldade enfrentada foi a aquisição de dados de expressão gênica que disponibilizassem informações clínicas, de sobrevida associados com número amostral consideravelmente robusto. Por fim, outro fator importante foi a coleta e criação de um banco de amostras de tumores pancreáticos, uma vez que este tipo tumoral é considerado raro, e ao diagnóstico, os pacientes com ADP já apresentam estágio avançado de desenvolvimento tumoral, sem opções cirúrgicas para tratamento.

Capítulo VII

Conclusões

Este trabalho associou, pela primeira vez, a expressão de TULP3 à carcinogênese gástrica através de análises de bioinformática utilizando dados de expressão gênica disponíveis em bancos de dados públicos e também através de métodos imunohistoquímicos. Como estudo preliminar, usando dados de expressão gênica, os níveis elevados de *TULP3* foram associados a um prognóstico desfavorável para o tipo difuso de câncer gástrico. Entretanto quando realizamos a imunomarcação desta proteína, a positividade de TULP3 não apresentou associação a qualquer histotipo, devido ao baixo número de amostras com coloração positiva. Ao analisarmos o perfil de expressão gênica e proteica de TULP3 em lesões gástricas pré-malignas como gastrite crônica e metaplasia intestinal, identificamos um aumento de expressão gênica no subtipo incompleto de metaplasia. Ainda, ao verificar o envolvimento proteico de TULP3 nos tecidos gástricos, foi identificada uma associação entre a positividade da imunomarcação da proteína com as amostras de metaplasia.

Enquanto que para o adenocarcinoma colorretal apenas a expressão gênica de *TULP3* foi associada a preditores prognósticos como invasão vascular para o adenocarcinoma retal e invasão linfática para o adenocarcinoma de cólon. Esses achados preliminares indicam um possível papel de TULP3 na carcinogênese de tumores gástricos e colorretais, entretanto mais estudos são necessários para a definição de TULP3 como biomarcador para esses tipos tumorais.

Na avaliação do perfil de expressão proteica de TULP3 em tecidos pancreáticos, observamos uma associação entre o escore de imunorreatividade (IRS) de TULP3 positivo com tecido pancreático não-tumoral adjacente e lesões intraepiteliais pancreáticas (PanIN), e ainda, uma associação negativa com adenocarcinoma ductal pancreático. Esses achados indicam que TULP3 pode não estar associado à carcinogênese pancreática, mas no desenvolvimento celular de tecido pancreático não-tumoral adjacente e das lesões intraepiteliais pancreáticas.

Capítulo VIII

Perspectivas

Como perspectivas à continuidade desse trabalho, podemos mencionar:

- A análise do RNAm de *TULP3* dos casos de adenocarcinoma pancreático, tecido não-tumoral adjacente, tecidos ADP e de carcinoma de ampola de Vater, e sua correlação com o desfecho clínico dos pacientes.
- A análise da expressão proteica de TULP3 nas linhagens celulares de ADP (PANC-1, PANC-2 e CAPAN-1).
- A realização de análises de bioinformática correlacionando a expressão de *TULP3* com genes da via de sinalização Wnt / β -catenina para os adenocarcinomas pancreático, gástrico e colorretais e a posterior análise através da imunohistoquímica.
- A análise imunohistoquímica de TULP3 em tumores colorretais e lesões pré-malignas;
- A realização de estudos utilizando banco de dados público CPTAC-TCGA que contém informações do proteoma de adenocarcinoma colorretal, correlacionando os valores da expressão gênica de *TULP3* com a quantificação proteica.

Anexos

Colaboração em artigos durante o período de Doutorado relacionados aos seguintes temas:

Preliminary screening of histone deacetylases in carcinoma of the ampulla of Vater.

Cleandra Gregório^{1,2}; Ivaine Taís Sauthier Sartor^{1,2}; Bárbara Alemar^{1,2}; Raquel Camara Rivero^{3,5}; Simone Marcia dos Santos Machado⁵; Diego de Mendonça Uchoa⁶; Alessandro Bersch Osvaldt^{4,5}; Patricia Ashton-Prolla^{1,2,5}.

¹Programa de Pós-graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil.

²Laboratório de Medicina Genômica, Centro de Pesquisa Experimental - Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul, Brazil.

³Departamento de Patologia, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

⁴Grupo de Vias Biliares e Pâncreas – Cirurgia do Aparelho Digestivo, Hospital de Clínicas de Porto Alegre, Porto Alegre, Rio Grande do Sul, Brazil

⁵Hospital de Clínicas de Porto Alegre, Porto Alegre, Rio Grande do Sul, Brazil

Corresponding author: Cleandra Gregório

E-mail address: cleandra.gregorio@gmail.com

Full postal address: Hospital de Clínicas de Porto Alegre. Laboratório de Medicina Genômica. 2350, Ramiro Barcelos St. Porto Alegre/ Brazil- 90035-903.

Manuscrito a ser submetido em revista a definir.

**Inborn errors of metabolism and sudden unexpected death in infancy: a
panorama of the state of Rio Grande do Sul, Brazil**

de Bitencourt, F.H.¹; Sartor, I.T.S.^{1,2} Schwartz, I.V.D.^{3,4} ; Vianna, F.S.L.^{1,5}

¹Graduate Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.

²Genomic Medicine Laboratory, Experimental Research Service, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil.

³Department of Genetics, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.

⁴Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil.

⁵Laboratório de Pesquisa em Bioética e Ética na Ciência (LAPEBEC), Experimental Research Service, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil.

Corresponding/pre-publication author:

Ida Vanessa Doederlein Schwartz,

MD, PhD Medical Genetics Service

Manuscrito a ser submetido em revista a definir.

Effects and mechanism of Rapamicin treatment on Tuberous Sclerosis Complex fibroblasts

Clévia Rosset^{1,2}, Mariane Jaeger³, Eduardo C. F. Chiela⁴, Larissa Brussa Reis^{1,2}, Ivaine Taís Sauthier Sartor^{1,2}, Cristina Brinckmann Oliveira Netto⁵, Caroline Brunetto de Farias^{3,6}, Rafael Roesler^{3,7}, and Patricia Ashton-Prolla^{1,2,5,8}

¹Laboratório de Medicina Genômica – Centro de Pesquisa Experimental – Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul, Brazil.

²Programa de pós-graduação em genética e biologia molecular, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil.

³Laboratório de Câncer e Neurobiologia – Centro de Pesquisa Experimental – Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul, Brazil.

⁴Programa de pós-graduação em gastroenterologia e hepatologia da Famed, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil.

⁵Serviço de Genética Médica, HCPA, Porto Alegre, Rio Grande do Sul, Brazil.

⁶Instituto do Câncer Infantil, Porto Alegre, Rio Grande do Sul, Brazil.

⁷Departamento de Farmacologia - ICBS - UFRGS, Porto Alegre, Rio Grande do Sul, Brazil.

⁸Departamento de Genética - UFRGS, Porto Alegre, Rio Grande do Sul, Brazil.

(*) Correspondence to: Patricia Ashton Prolla. MD/ PhD

Serviço de Genética Médica. Hospital de Clínicas de Porto Alegre.

Rua Ramiro Barcelos 2350. CEP: 90035-903 - Porto Alegre - RS - Brazil.

e-mail: pprolla@hcpa.edu.br Tel/Fax: + 55 51 3359 8011

Manuscrito submetido à Scientific Reports.

Clinical and molecular characterization of patients fulfilling Chompret criteria for Li-Fraumeni Syndrome in Southern Brazil

Camila Matzenbacher Bittar^{a,b}, Igor Araujo Vieira^{a,b}, Clévia Rosset^{a,b}, Tiago Finger Andreisa,^b Yasminne de Oliveira Roch^{a,b}, Ivaine Tais Sauthier Sartor^{a,b,c}, Osvaldo Artigalás^c, Cristina B. O. Netto^d, Barbara Alemar^{a,b}, Gabriel S. Macedo^b, Patricia Ashton-Prolla^{a,b,d,**}

a Programa de Pós-Graduação em Genética e Biologia Molecular (PPGBM), Universidade Federal do Rio Grande do Sul (UFRGS), Av. Bento Gonçalves, 9500 - Prédio 43323M, Porto Alegre, Rio Grande do Sul 91501-970, Brazil;

b Laboratório de Medicina Genômica, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre (HCPA), Rua Ramiro Barcelos, 2350, Porto Alegre, Rio Grande do Sul 90035-903, Brazil;

c Hospital Moinhos de Vento (HMV), Rua Ramiro Barcelos, 910, Porto Alegre, Rio Grande do Sul 91790-560, Brazil;

d Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre (HCPA), Rua Ramiro Barcelos, 2350, Porto Alegre, Rio Grande do Sul 90035-903, Brazil.

**Corresponding author: Camila Matzenbacher Bittar

E-mail address: camilambittar@gmail.com

Telephone number: + 55 51 3333-7933

Manuscrito a ser submetido em revista a definir

Referências

Allgayer H, Babic R, Gruetzner KU, Tarabichi A, Schildberg FW, Heiss MM. c-erbB-2 is of independent prognostic relevance in gastric cancer and is associated with the expression of tumor-associated protease systems. *J Clin Oncol*. 2000;18:2201-9

Al-Shamsi HO, Jones J, Fahmawi Y, Dahbour I, Tabash A, Abdel-Wahab R, et al. Molecular spectrum of KRAS, NRAS, BRAF, PIK3CA, TP53 and APC somatic gene mutations in Arab patients with colorectal cancer: determination of frequency and distribution pattern. *J Gastrointest Oncol*. 2016;7(6):882-902

Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin*. 2017;67:93-99

aPark IJ, Choi GS, Lim KH, Kang BM, Jun SH. Serum carcinoembryonic antigen monitoring after curative resection for colorectal cancer: clinical significance of the preoperative level. *Ann Surg Oncol*. 2009;16:3087-3093

Aran D, Camarda R, Odegaard J, Paik H, Oskotsky B, Krings G, Goga A, et al. Comprehensive analysis of normal adjacent to tumor transcriptomes. *Nat Commun*. 2017;8(1):1077

Ardengh JC, Coelho N, Osvaldt AB. Câncer de pâncreas em fase inicial: é possível identifica-lo através dos instrumentos científicos e propedêuticos atualmente disponíveis? *Arquivos de Gastroenterologia*. 45:168-177

Assumpção PP, Santos SEB, Santos AKR, Demachki S, Khayat AS, Ishak G, et al. The adjacent tumor sample trap. *Gastric Cancer*. 2016;19(3):1024-1025

Bai YQ, Yamamoto H, Akiyama Y, Tanaka H, Takizawa T, Koike M, et al. Ectopic expression of homeodomain protein CDX2 in intestinal metaplasia and carcinomas of the stomach. *Cancer Lett*. 2002;176(1):47-55

Ballehaninna UK & Chamberlain RS. Biomarkers for pancreatic cancer: Promising new markers and options beyond CA19-9. *Tumour Biol.* 2013;34:3279-3292

Ballester V, Rashtak S, Boardman L. Clinical and molecular features of young-onset colorectal cancer. *World J Gastroenterol.* 2016;22(5):1736-1744

Balmana J, Balaguer F, Cervantes A, Arnold D, Group EGW. Familial risk-colorectal cancer: ESMO Clinical Practice Guidelines. *Annals of oncology: official journal of the European Society for Medical Oncology / ESMO.* 2013;24 Suppl 6:vi73-80.

Bang YJ, Cutsem EV, Feyereislova A, Chung HC, Shen L, Sawaki A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2- positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomized controlled trial. *Lancet.* 2010;376:687-697

Baniak N, Senger JL, Ahmed S, Kanthan SC, Kanthan R. Gastric biomarkers: a global review. *World J Surg Oncol.* 2016;14:212

Bardeesy N & DePinho RA. Pancreatic cancer biology and genetics. *Nat Rev Cancer.* 2002;2(12):897-909

Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for functional genomics datasets - 10 years on. *Nucleic Acids Res.* 2011;39(Database issue):D1005-D1010

Begnami MD, Fukuda E, Fregnani JH, Nokogaki S, Montagnini AL, da Costa WLJ, et al. Prognostic implications of altered human epidermal growth factor receptors (HERs) in gastric carcinomas: HER2 and HER3 are predictors of poor outcome. *J Clinical Oncol.* 2011;29(22):3030-26

Bhargava R, Beriwal S, McManus K, Dabbs DJ. CK5 is more sensitive than CK5/6 in identifying the "basal-like" phenotype of breast carcinoma. *Am. J. Clin. Pathol.* 2008;130(5):724-730

Bilimoria KY, Bentrem DJ, Ko CY, Ritchey J, Stewart AK, Winchester DP, et al. Validation of the 6th Edition AJCC Pancreatic Cancer Staging System. *Cancer.* 2007;110:738-744. American Cancer Society

Bilous M, Osamura RY, Rüschoff J, et al. HER-2 amplification is highly homogenous in gastric cancer. *Hum Pathol* 2010; 41:304-5; author reply 305-6

Boggon TJ, Shan WS, Santagata S, Myers SC, Shapiro L. Implication of tubby proteins as transcription factors by structure-based functional analysis. *Science*. 1999;286:2119-2125

Boku N. HER2-positive gastric cancer. *Gastric cancer*. 2014;17(1):1-12

Bond-Smith G, Banga N, Hammond TM, Imber CJ. Pancreatic adenocarcinoma. *BMJ*. 2012;344:e2476

Bornschein J, Kandulski A, Selgrad M, Malfertheiner P. From gastric inflammation to gastric cancer. *Dig Dis*. 2010;28:609-614

Bosman FT, Carneiro F, Hruban RH, Theise ND, editors. WHO classification of tumors. Lyon: WHO Press; 2010:292-295

bPark IJ, Choi GS, Jun SH. Prognostic value of serum tumor antigen CA19-9 after curative resection of colorectal cancer. *Anticancer Res*. 2009;29:4303-4308

Brooks JD. Translational genomics: the challenge of developing cancer biomarkers. *Genome Res*. 2012;22(2):183-187

Budinska E, Popovici V, Tejpar S, D'Ario G, Lapique N, Sikora KO, et al. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *J. Pathol*. 2013;231:63-76

Calistri D, Rengucci C, Seymour I, et al. Mutation analysis of TP53, K-ras, and BRAF genes in colorectal cancer progression. *J Cell Physiol*. 2005;204:484-8

Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*. 2014;513:202-209

Cancer Genome Atlas Research Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487:330-337

Carro MS, Lim WK, Alvarez MK, Bollo RJ, Zhao X, Snyder EY, et al. The transcriptional network for mesenchymal transformation of brain tumours. *Nature*. 2010;463(21):318-325

Chang JC & Kundranda M. Novel diagnostic and predictive biomarkers in pancreatic adenocarcinoma. *Int J Mol Sci*. 2017;18(3):E667

Cho SY, Park JW, Liu Y, Park SY, Kim JH, Yang H, et al. Sporadic Early-onset diffuse gastric cancers have high frequency of somatic CDH1 alterations, but low frequency of somatic RHOA mutations compared with late-onset cancers. *Gastroenterology*. 2017;153:536-549.e26

Chong Y, Mia-Jan K, Ryu H, Abdul-Ghafar J, Munkhdelger J, Lkhagvadorj S, et al., DNA methylation status of a distinctively different subset of genes is associated with each histologic Laurén classification subtype in early gastric carcinogenesis. *Oncology reports*. 2014;31(6):2535-2544

Coleman DL & Eicher EM. Fat (fat) and Tubby (tub): two autosomal recessive mutation causing obesity syndromes in the mouse. *Journal of Heredity*. 1990;81:424-427

Companioni O, Sanz-Anquela JM, Pardo ML, Puigdecenet E, Nonell L, García N, et al. Gene expression study and pathway analysis of histological subtypes of intestinal metaplasia that progress to gastric cancer. *PLoS One* 2017;12(4):e0176043

Correa P & Piazzuelo MB. The gastric precancerous cascade. *J Dig Dis*. 2012;13(1):2-9

Correa P. Gastric cancer: overview. *Gastroenterol Clin North Am*. 2013;42(2):211-217

Costa NR, Gil da Costa RM, Medeiros R. A viral map of gastrointestinal cancers. *Life Sci*. 2018;199:188-200

Coverlizza S, Risio M, Ferrari A, Fenoglio-Preiser CM, Rossini FP. Colorectal adenomas containing invasive carcinoma: pathologic assessment of lymph node metastatic potential. *Cancer*. 1989;64(9):1937-1947.

Cristescu R, Lee J, Nebozhyn M, Kim Kyoung-Mee, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med*. 2015;21:449-456

Cutsem ER, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. *Lancet*. 2016;388:2654-2664

De Sousa E Melo F, Wang X, Jansen M, Fessler E, Trinh A, de Rooij LP, et al. Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nat. Med.* 2013;19:614-618

Decker GA, Batheja MJ, Collins JM, Silva AC, Mekeel KL, et al. Risk factors for pancreatic adenocarcinoma and prospects for screening. *Gastroenterol Hepatol.* 2010;6(4):246-254

DeVita VTLT & Rosenberg SA *Cancer: Principles & Practice of Oncology*. Edition t, editor. Philadelphia: Lippincott Williams & Wilkins. 2014.

Dienstmann R, Guinney J, Delorenzi M, Reynies AD, Roepman P, Sadanandam A et al. Colorectal Cancer Subtyping Consortium (CRCSC) identification of a consensus of molecular subtypes. *Journal of Clinical Oncology*. 2014;32(15_suppl):3511-3511.

Dikken JL, van de Velde CJ, Coit DG, Shah MA, Verheij M, Cats A. Treatment of resectable gastric cancer. *Therap Adv Gastroenterol.* 2012;5(1):49-69

Eda A, Osawa H, Yanaka I, Satoh K, Mutoh H, Kihira K, et al. Expression of homeobox gene CDX2 precedes that of CDX1 during the progression of intestinal metaplasia. *J Gastroenterol.* 2002;37:94-100

Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002;30:207-210

Fausett SR & Klingensmith J. Compartmentalization of the foregut tube: developmental origins of the trachea and esophagus. *Dev Biol.* 2012;1:184-202

Fearon ER & Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell.* 1990;61(5):759-767.

Fenoglio-Preiser C, Carneiro F, Correa P, Guilford P, Lambert R, Megraud F, et al. Tumours of the stomach. In Halmilton SR, Aaltonen LA, editors. *World Health Organization Classification of*

Tumours. Pathology and Genetics of Tumours of the Digestive System. IARC Press: Lyon 2000. p. 37-68.

Fernandes-Zapico ME, Bramati PS, Zakaria S, Kaczynski JA, Urrutia R. Fundamentals of transcription factors and their impact on pancreatic development and cancer. *Pancreatology*. 2003;3:276-283

Fraum TJ, Owen JW, Flower KJ. Beyond histologic staging: emerging imaging strategies on colorectal cancer with special focus on magnetic resonance imaging. *Clin Colon Rectal Surg*. 2016;29(3):205-215

Fujii T, Sutoh T, Morita H, Yajima R, Yamaguchi S, Tsutsumi S, et al. Vascular invasion, but not lymphatic invasion, of the primary tumor is a strong prognostic factor in patients with colorectal cancer. *Anticancer Res*. 2014;34(6):3147-3152.

Garcia I, Vizoso F, Martín A, Sanz L, Abdel-Lah O, Raigoso P, et al. Clinical significance of the epidermal growth factor receptor and HER2 receptor in resectable gastric cancer. *Ann Surg Oncol*. 2003;10:234-41

Ge S, Xia X, Ding C, Zhen B, Zhou Q, Feng J, et al. A proteomic landscape of diffuse-type gastric cancer. *Nat Commun*. 2018;9:1012

Glimelius B, Tiret E, Cervantes A, Arnold D, Group EGW. Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of oncology: official journal of the European Society for Medical Oncology / ESMO*. 2013;24 Suppl 6:vi81-8.

Globocan 2012 (IARC). Estimated cancer incidence, mortality, and prevalence worldwide in 2012. From: <http://globocan.iarc.fr/old/FactSheets/cancers.asp> Acessado em 03 de Janeiro de 2019

Goldberg EM, Simunovic LM, Drake SL, Mueller WFJ, Verrill HL. Comparison of serum CA 19–9 and CEA levels in a population at high risk for colorectal cancer. *Hybridoma*. 1989;8:569-575

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*. 1999;286:531-537.

Gonzaga IS, Lima, SCS, Nicolau MC, Nicolau-Neto P, Costa NM, Simão TA, et al. TFF1 hypermethylation and decreased expression. In esophageal squamous cell carcinoma and histologically normal tumor surrounding esophageal cells. 2017;9:130

González CA, Sanz-Anquela JM, Companioni O, Bonet C, Berdasco M, Lopez C, et al. Incomplete type of intestinal metaplasia has the highest risk to progress to gastric cancer: results of the Spanish follow-up multicenter study. *J Gastroenterol Hepatol*. 2016;31(5):953-958

González CA, Sanz-Anquela JM, Gisbert JP, Correa P. Utility of subtyping intestinal metaplasia as marker of gastric cancer risk. A review of evidence. *Int J Cancer*. 2013;133(5):1023-1032

Han T, Lin J, Wang Y, Fan Q, Sun H, Tao Y, et al. Forkhead box D1 promotes proliferation and suppresses apoptosis via regulating polo-like kinase 2 in colorectal cancer. *Biomed Pharmacother*. 2018;103:1369-1375

Harada T, Chelala C, Crnogorac-Jurcevic T, Lemoine NR. Genome-wide analysis of pancreatic cancer using microarray-based techniques. *Pancreatol*. 2009;9:13-24

Harris EI, Lewin DN, Wang HL, Lauwers GY, Srivastava A, Shyr Y, et al. Lymphovascular invasion in colorectal cancer: an interobserver variability study. *Am J Surg Pathol*. 2008;32(12):1816-1821

Harsha HC, Kandasamy K, Ranagathan P, Rani S, Ramabadran S, Gollapudi S, et al. A compendium of potential biomarkers of pancreatic cancer. *Plos Med*. 2009;6(4):e1000046

Harsha HC, Kandasamy K, Ranagathan P, Rani S, Ramabadran S, Gollapudi S, et al. A compendium of potential biomarkers of pancreatic cancer. *Plos Med*. 2009;6(4):e1000046

Herreros-Villanueva M, Hijona E, Cosme A, Bujanda L. Adjuvant and neoadjuvant treatment in pancreatic cancer. *World J Gastroenterol*. 2012;18(14):1565-1572

Hezel AF, Kimmelman AC, Stanger BZ. Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev*. 2006;20:1218-1249

Hidalgo M. Pancreatic cancer. *N Engl J Med*. 2010;362(17):1605-1617

Hofmann M, Stoss O, Shi D, Buttner R, van de Vijver M, Kim W, et al. Assessment of a HER2 scoring system for gastric cancer: results from a validation study. *Histopathology*. 2008;52:797-805

Hruban RH, Adsay NV, Albores-Saavedra J, Compton C, Garrett ES, Goodman SN, et al. Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions. *Am J Surg Pathol*. 2001;25(5):579-586

Hu B, Hajj NE, Sittler S, Lammert N, Barnes R, Meloni-Ehrig A. Gastric cancer: Classification, histology and application of molecular pathology. *J Gastrointest Oncol*. 2012;3(3):251-261

INCA. Câncer. Tipos de câncer.
<http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/> Acessado em 03 Janeiro de 2019

Jennings RE, Berry AA, Strutt JP, Gerrard DT, Hanley NA. Human pancreas development. *Devel*. 2015;142:3126-3137

Kang JM, Lee BH, Kim N, Lee HS, Lee HE, Park JH, et al. CDX1 and CDX2 Expression in Intestinal Metaplasia, Dysplasia and Gastric Cancer. *J Korean Med Sci*. 2011;26:647-653

Kang JM, Lee BH, Kim N, Lee HS, Lee HE, Park JH, et al. CDX1 and CDX2 expression in intestinal metaplasia, dysplasia and gastric cancer. *J Korean Med Sci*. 2011;26(5):647-653

Kawakami H, Zaanan A, Sinicrope FA. Microsatellite instability testing and its role in the management of colorectal cancer. *Curr Treat Options Oncol*. 2015;16(7):30

Khan MA, Azim S, Zubair H, Bhardwaj A, Patel GK, Khushman M, et al. Molecular drivers of pancreatic cancer pathogenesis: looking inward to move forward. *Int J Mol Sci*. 2017;18(4):779

Kim MJ. The usefulness of CDX-2 for differentiating primary and metastatic ovarian carcinoma: an immunohistochemical study using a tissue microarray. *J Korean Med Sci*. 2005;20(4):643-648

Kim TH, Shivdasani RA. Stomach development, stem cells and disease. *Devel*. 2016;143:554-565

Kleyn PW, Fan W, Kovats SG, Lee JJ, Pulido JC, Wu Y, et al. Identification and characterization of the mouse obesity gene *tubby*: a member of a novel gene family. *Cell*. 1996;85:281-290

Koumariou A, Tzeveleki I, Mekras D, Eleftheraki AG, Bobos M, Wirtz R, et al. Prognostic markers in early-stage colorectal cancer: significance of TYMS mRNA expression. *Anticancer Res*. 2014;34(9):4949-4962.

Kovacheva VN & Rajpoot NM. Subcellular protein expression models for microsatellite instability in colorectal adenocarcinoma tissue images. *BMC Bioinformatics*. 2016;17:430

Labianca R, Nordlinger B, Beretta GD, Mosconi S, Mandalà M, Cervantes A, et al. Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of oncology: official journal of the European Society for Medical Oncology / ESMO*. 2013;24 Suppl 6:vi64-72.

Laurén P. The two histological main types of gastric carcinoma. Diffuse and so-called intestinal type carcinoma: an attempt at histoclinical classification. *Acta Pathol Microbiol Scand* 1965; 64:31-49.

Li C, Shen Z, Zhou Y, Yu W. Independent prognostic genes and mechanism investigation for colon cancer. *Biol Res*. 2018;51(1):10.

Li D, Xie K, Wolff R, Abbruzzese JL. Pancreatic cancer. *Lancet*. 2004;363:1049-1057.

Lin HH, Wei NC, Chou TY, Lin CC, Lan YT, Chang SC, et al. Building personalized treatment plans for early-stage colorectal cancer patients. *Oncotarget*. 2017;8(8):13805-13817

Lindor NM, Burgart LJ, Leontovich O, Goldberg RM, Cunningham JM, Sargent DJ, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J Clin Oncol*. 2002;20:1043-1048

Liu L, Wang ZW, Ji J, Zhang JN, Yan M, Zhang J, et al. A cohort study and meta-analysis between histopathological classification and prognosis of gastric carcinoma. *AntiCancer Agents Med Chem*. 2013;13(2):227-234

Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer*. 2005;5(11):845-56

Maitra A & Hruban RH. Pancreatic cancer. *Annu Ver Pathol*, 2008;157-188

Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera RD, et al. ARACNe: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*. 2006;7:S7

Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med*. 2013;10:e1001453

Matsuoka T & Yashiro M. Biomarkers of gastric cancer: current topics and future perspective. *World J Gastroenterol*. 2018;24(26):2818-2832

Mayo SC, Austin DF, Sheppard BC, Mori M, Shipley DK, Billingsley KG. Evolving preoperative evaluation of patients with pancreatic cancer: does laparoscopy have a role in the current era? *J Am Coll Surg*. 2009;208(1):87-95

McLean MH & El-Omar EM. Genetics of gastric cancer. *Nat Rev Gastroenterol Hepatol*. 2014;11(11):664-674

Mehta S, Shelling A, Muthukaruppan A, Lasham A, Blenkiron C, Laking G, et al. Predictive and prognostic molecular markers for cancer medicine. *Ther Adv Med Oncol*. 2010;2:125-148

Mizoshita T, Inada K, Tsukamoto T, Kodera Y, Yamamura Y, Hirai T. Expression of Cdx1 and Cdx2 mRNAs and relevance of this expression to differentiation in human gastrointestinal mucosa--with special emphasis on participation in intestinal metaplasia of the human stomach. *Gastric Cancer*. 2001;4(4):185-191

Moasser, MM. The oncogene HER2; Its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene*. 2007;26(45): 6469-6487

Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, et al. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487:330-337

- Mukhopadhyay S & Jackson PK. The tubby family proteins. *Genome Biology*. 2011;12:22
- Mukhopadhyay S, Wen X, Chih B, Nelson CD, Lane WS, Scales SJ, et al. TULP3 bridges the IFT-A complex and membrane phosphoinositides to promote trafficking of G protein-coupled receptors into primary cilia. *Genes Dev*. 2010;24:2180-2193
- Mukhopadhyay S, Wen X, Chih B, Nelson CD, Lane WS, Scales SJ, et al. TULP3 bridges the IFT-A complex and membrane phosphoinositides to promote trafficking of G protein-coupled receptors into primary cilia. *Genes Dev*. 2010;24:2180-2193
- Muller S, Chesner IM, Egan MJ, Rowlands DC, Collard MJ, Swarbrick ET, et al. Significance of venous and lymphatic invasion in malignant polyps of the colon and rectum. *Gut*. 1989;30(10):1385-1391
- Nagini S. Carcinoma of the stomach: a review of epidemiology, pathogenesis, molecular genetics and chemoprevention. *World J Gastrointest Oncol*. 2012;4(7):156-169
- Nakagawa N, Li J, Yabuno-Nakagawa K, Eom TY, Cowles M, Mapp T, et al. APC sets the Wnt tone necessary for cerebral cortical progenitor development. *Genes Dev*. 2017;31(16):1679-1692
- Nalejska E, Maczynska E, Lewandowska MA. Prognostic and predictive biomarker: tools in personalized oncology. *Mol Diagn Ther*. 2014;18(3):273-284
- Nishina PM, North MA, Ikeda A, Yan Y, Naggert JK. Molecular characterization of a novel tubby gene family member, TULP3 in mouse and humans. *Genomics*. 1998;54(2):215-220.
- Nitsche C, Simon P, Weiss FU, Fluhr G, Weeber E, Gärtner S, et al. Environmental risk factors for chronic pancreatitis and pancreatic cancer. *Dig Dis*. 2011;29(2):235-42
- Nivatvongs S, Rojanasakul A, Reiman HM, Dozois RR, Wolff BG, Pemberton JH, et al. The risk of lymph node metastasis and colorectal polyps with invasive carcinoma. *Dis Colon Rectum*. 1991;34(4):323-328
- Norman RX, Ko HW, Huang V, Eun CM, Ablner LL, Zhang Z, et al. Tubby-like protein 3 (TULP3) regulates patterning in the mouse embryo through inhibition of Hedgehog signaling. *Hum Mol Genet*. 2009;18(10):1740-1754

Ntzani EE & Ioannidis JP. Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment. *Lancet*. 2003;362(9394):1439-1444

Ohlemiller KK, Hughes RM, Mosinger-Ogilvie J, Speck JD, Grosf DH, Silverman MS. Cochlear and retinal degeneration in the tubby mouse. *Neuroreport*. 1995;6:845-849

Ouchi K, Sugawara T, Ono H, Fujiya T, Kamiyama Y, Kakugawa Y, et al. Histologic features and clinical significance of venous invasion in colorectal carcinoma with hepatic metastasis. *Cancer*. 1996;78(11):2313-2317

Payne SJL, Bowen RL, Jones JL, Wells CA. Predictive markers in breast cancer – the present. *Histopathology*. 2008;52:82-90

Perez-Villamil B, Romera-Lopez A, Hernandez-Prieto S, Lopez-Campos G, Calles A, Asenjo JAL, et al. Colon cancer molecular subtypes identified by expression profiling and associated to stroma, mucinous type and different clinical behavior. *BMC Cancer*. 2012;12:260

Peyssonnaud C & Eychene A. The Raf/MEK/ERK pathway: new concepts of activation. *Biol Cell* 2001;93:53-62

Pinchuk IV, Mifflin RC, Saada JI, Powell DW. Intestinal Mesenchymal Cells. *Curr Gastroenterol Rep*. 2010;12:310-318

Qian X, Tan H, Zhang J, Liu K, Yang T, Wang M, et al. Identification of biomarkers for pseudo and true progression of GBM based on radiogenomics study. *Oncotarget*. 2016;7(34):55377-55394

R Development Core Team (2011) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Viena, Austria. ISBN 3-900051-070 <http://www.r-project.org/>

Roepman P, Schlicker A, Tabernero J, Majewski I, Tian S, Moreno V, et al. Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition. *Int. J. Cancer*. 2014;134:552-562

Rustici G, Kolesnikov N, Brandizi M, Burdett T, Dylag M, Emam I, et al. ArrayExpress update-trends in database growth and links to data analysis tools. *Nucleic Acids Res.* 2013;41:D987–D990

Ryan DP, Hong TS, Bardeesy N. Pancreatic Adenocarcinoma. *N Engl J Med.* 2014;371:1039-1049

Sadanandam A, Lyssiotis CA, Homicsko K, Collisson EA, Gibb WJ, Wullschleger S, et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat. Med.* 2013;19:619-625

Santagata S, Boggon TJ, Baird CL, Gomez CA, Zhao J, Shan WS, et al. G-protein signaling through tubby proteins. *Science.* 2001;292:2041-2050

Sanz-Pamplona R, Cordero D, Berenguer A, Lejbkowitz F, Rennert H, Salazar R, et al. Gene expression differences between colon and rectum tumors. *Clin Cancer Res.* 2011;17(23):7303-7312

Sartor ITS, Zeidán-Chuliá F, Albanus RD, Dalmolin RJS, Moreira JCF. Computational Analyses Reveal a Prognostic Impact of TULP3 as a Transcriptional Master Regulator in Pancreatic Ductal Adenocarcinoma. *Mol Biosyst.* 2014;10:1461-1468

Schlicker A, Beran G, Chresta CM, McWalter G, Pritchard A, Weston S, et al. Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. *BMC Med. Genomics.* 2012;5:66

Schmidt N, Peitz U, Lippert H, Malfertheiner P. Missing gastric cancer in dyspepsia. *Aliment Pharmacol Ther.* 2005;21(7):813-820

Seno H, Oshima M, Taniguchi MA, Usami K, Ishikawa TO, Chiba T, et al. CDX2 expression in the stomach with intestinal metaplasia and intestinal-type cancer: prognostic implications. *Int J Oncol* 2002;21(4):769-74

Shang J & Pena AS. Multidisciplinary approach to understand the pathogenesis of gastric cancer. *World J Gastroenterol.* 2005;11(27):4131-4139

Sharma K, Bhatia V, Agarwal P, Kumar A. Gastrointestinal cancers: molecular genetics and biomarkers. *Can J Gastroenterol Hepatol*. 2018;45:13860

Shimada H, Noie T, Ohashi M, Oba K, Takahashi Y. Clinical significance of serum tumor markers for gastric cancer: a systematic review of literature by the Task Force of the Japanese Gastric Cancer Association. *Gastric Cancer*. 2014;17:26-33

Shin CM, Kim N, Chang H, Kim JS, Lee DH, Jung HC. Follow-Up Study on CDX1 and CDX2 mRNA expression in noncancerous gastric mucosae after *Helicobacter pylori* eradication. *Dig Dis Sci*. 2016;61(4):1051-1059

Shinohara H, Morita S, Kawai M, Miyamoto A, Sonoda T, Pastan I, et al. Expression of HER2 in human gastric cancer cells directly correlates with antitumor activity of a recombinant disulfide-stabilized anti-HER2 immunotoxin. *J Surg Res*. 2002;102:169-77

Siegel R, Naishadham D, Jemal A. Cancer statistics. *CA Cancer J Clin*. 2012;62(1):10-29

Sobin LH & Wittekind CH. TNM: classificação de tumores malignos. 6ed. INCa, Rio de Janeiro. 2004;254

Sotiriou C & Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nat Rev Cancer*. 2007;7:545-553

Staab HJ, Brummendorf T, Hornung A, Anderer FA, Kieninger G. The clinical validity of circulating tumor-associated antigens CEA and CA 19–9 in primary diagnosis and follow-up of patients with gastrointestinal malignancies. *Klin Wochenschr*. 1985;63:106-115

Takahashi T, Saikawa Y, Kitagawa Y. Gastric cancer: current status of diagnosis and treatment. *Cancers (Basel)*. 2013;5(1):48-63

Takehana T, Kunitomo K, Kono K, Kitahara F, Iizuka H, Matsumoto Y, et al. Status of c-erbB-2 in gastric adenocarcinoma: a comparative study of immunohistochemistry, fluorescence in situ hybridization and enzyme-linked immuno-sorbent assay. *Int J Cancer*. 2002; 98:833-7

Tam PPL, Kanai-Azumay M, Kanaiz Y. Early endoderm development in vertebrates: lineage differentiation and morphogenetic function. *Curr Opin Genetics Dev*. 2003;13:393-400

Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science*. 1993;260:816-819

Umbreit C, Flanjak J, Weiss C, Erben P, Aderhold C, Faber A, et al. Incomplete Epithelial–Mesenchymal Transition in p-16 positive squamous cell carcinoma cells correlates with β -catenin expression. *Anticancer Res*. 2014;34(12):7061-7070

van't Veer LJ & Bernards R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature*. 2008;452:564-570

Vincent A, Herman J, Schulick R, Hruban RH, Goggins M. Pancreatic cancer. *Lancet*. 2011;378:607-620

Wan Q, Dingerdissen H, Fan Y, Gulzar N, Pan Y, Wu TS. BioXpress: an integrated RNA-seq-derived gene expression database for pan-cancer analysis. *Database*. 2015;2015

Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet*. 2005;365(9460):671-679

Xing Y, Zhao Z, Zhu Y, Zhao L, Zhu A, Piao D. Comprehensive analysis of differential expression profiles of mRNAs and lncRNAs and identification of a 14-lncRNA prognostic signature for patients with colon adenocarcinoma. *Oncol Rep*. 2018;39(5): 2365-2375

Yakabe T, Nakafusa Y, Sumi K, Miyoshi A, Kitajima Y, Sato S, et al. Clinical significance of CEA and CA19-9 in postoperative follow-up of colorectal cancer. *Ann Surg Oncol*. 2010;17:2349-2356

Yokoyama Y, Nimura Y, Nagino M. Advances in the treatment of pancreatic cancer: limitations of surgery and evaluation of new therapeutic strategies. *Surgery Today*. 2009;39(6):466-475

Yoshida N, Kinugasa T, Ohshima K, Yuge K, Ohchi T, Fujino S, et al. Analysis of Wnt and β -catenin Expression in Advanced Colorectal Cancer. *Anticancer Res*. 2015;35(8):4403-4410

Yu Z, Chen Z, Wu J, Li Z, Wu Y. Prognostic value of pretreatment serum carbohydrate antigen 19-9 level in patients with colorectal cancer: A meta-analysis. *PLoS One*. 2017;12(11):e0188139

Zeng Z, Cohen AM, Urmacher C. Usefulness of carcinoembryonic antigen monitoring despite normal preoperative values in node-positive colon cancer patients. *Dis Colon Rectum*. 1993;36:1063-1068