



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

Marco Antônio Marques Pretti

Metastização em tumores de cavidade oral HPV negativos: uma abordagem *in silico*

Orientador: Dr. Martín Hernan Bonamino

Co-orientadora: Dra. Mariana Lima Boroni Martins

Co-orientadora: Dra. Luciana Barros Carvalho

RIO DE JANEIRO
2019



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

Marco Antônio Marques Pretti

Metastização em tumores de cavidade oral HPV negativos: uma abordagem *in silico*

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

Orientador: Dr. Martin Hernan Bonamino
Co-orientadora: Dra. Mariana Lima Boroni Martins
Co-orientadora: Dra. Luciana Barros Carvalho

RIO DE JANEIRO
2019

P922m Pretti, Marco Antônio Marques.

Metastização em tumores de cavidade oral HPV negativos: uma abordagem *in silico* / Marco Antônio Marques Pretti. – Rio de Janeiro, 2019.
90 f. : il. color.

Dissertação (Mestrado em Oncologia) – Instituto Nacional de Câncer José Alencar Gomes da Silva, 2019.

Orientador: Martín Hernan Bonamino.

Co-orientadores: Mariana Lima Boroni Martins; Luciana Barros Carvalho.

1. Neoplasias Bucais. 2. Metástase Neoplásica. 3. Evasão Tumoral. I. Bonamino, Martín Hernan (Orient.). II. Martins, Mariana Lima Boroni (Co-orient.). III. Carvalho, Luciana Barros (Co-orient.). IV. Instituto Nacional de Câncer José Alencar Gomes da Silva. V. Título.

CDD 616.99431



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

Marco Antônio Marques Pretti

Metastização em tumores de cavidade oral HPV negativos: uma abordagem *in silico*

Orientador: Dr. Martin Hernan Bonamino

Co-orientadora: Dra. Mariana Lima Boroni Martins

Co-orientadora: Dra. Luciana Barros Carvalho

Aprovado em: 19/02/2019

Examinadores:

Dra. Sheila Coelho Soares Lima

Dr. Pedro Alexandre Favoretto Galante

Dra. Adriana César Bonomo

Dr. João Paulo de Biaso Viola

Dra. Tatiana de Almeida Simão

RIO DE JANEIRO
2019

Agradecimentos

Obrigado aos pacientes com câncer que se voluntariaram e doaram amostras biológicas para este e outros estudos.

Obrigado à Deus pelo dom da vida e à minha família pela educação e amor.

Agradeço aos meus orientadores Martín Bonamino, Mariana Boroni e Luciana Barros pela amizade, discussões científicas, ensinamentos e broncas; pelos anos que passaram e pelo que estão por vir.

Aos nossos colaboradores do AC Camargo, Emmanuel Dias-Neto, Frederico Netto, Diana Nunes, Israel Tojal e Jaqueline Ramalho um agradecimento especial pela confiança depositada em nós para juntos desenvolvermos este trabalho.

Obrigado aos amigos e colegas de pesquisa pelas conversas, ideias e apoio nos momentos difíceis. Em especial à Nicole Scherer, Carol Poubel, Cristóvão Lanna, Jéssica Cruz, Natasha Andressa, Leonardo Chicaybam, Luiza Abdo, Priscila Souza, Bárbara (Babi) Peixoto, Rômulo Galvani, Alice Sampaio e a todos os professores que incentivaram e colaboraram com minha carreira científica.

Agradeço aos desenvolvedores de programas e pacotes que auxiliam no avanço da bioinformática, Michael Love e muitos outros.

Por fim, agradeço à Mariana Muniz pela revisão do texto e por me aturar nessas semanas difíceis.

“...a plenitude da atividade humana é alcançada somente quando nela coincidem, se acumulam, se exaltam e se mesclam o trabalho, o estudo e o jogo; isto é, quando nós trabalhamos, aprendemos e nos divertimos, tudo ao mesmo tempo (...) é o que eu chamo de ócio criativo”

Domenico de Masi



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

Metastização em tumores de cavidade oral HPV negativos: uma abordagem *in silico*

Marco Antônio Marques Pretti

Resumo

Os tumores de cavidade oral são o quinto tumor mais frequente em homens conforme estimativa do INCa. A ocorrência de metástase durante o desenvolvimento do tumor aumenta as chances de recidiva e dificulta o tratamento. Tendo isso em vista, estudar os mecanismos que favorecem a metastatização desse tumor podem auxiliar na identificação de biomarcadores associados a este processo. Para tal, foi realizada uma análise *in silico* com enfoque imunológico de tumores de cavidade oral HPV negativos comparando tumores metastáticos e não-metastáticos. Foram analisadas 20 amostras de uma coorte em colaboração com o AC Camargo (ACC) e 22 amostras do repositório público TCGA (*The Cancer Genome Atlas*). Os dados brutos da coorte ACC foram pré-processados com avaliação da qualidade das leituras pelo FastQC e seleção das mesmas pelo Trimmomatic. O alinhamento contra o genoma humano GRCh38 foi realizado com o alinhador star e a expressão dos genes contabilizada com o pacote RSEM. Os genes diferencialmente expressos (DEG) foram avaliados pelo DESeq2, as vias enriquecidas pela plataforma Webgestalt e o microambiente tumoral com o xCell. Para os dados disponíveis no TCGA foi também identificado o alelo de HLA de cada amostra pelo Optitype para subsequente predição de neoantígenos com o netMHCpan. Os repertórios de receptores de linfócitos T (TCR) e B (BCR) foram identificados com o programa MiXCR. Resultado: Foram identificados 186 DEG para a coorte do TCGA, 127 para a coorte ACC e três em comum, sendo dois deles aumentado na mesma condição em ambas as coortes: PIWIL2 (mais expresso no grupo não-metastático) e ADH1B (mais expresso no grupo metastático). A população imunológica que mais correlacionou com o desfecho foi o linfócito T CD4 de memória (coeficiente de Pearson: -0,78 - ACC e -0,71 - TCGA) cuja assinatura está enriquecida no grupo não-metastático. Foi identificado um maior número de clones de TCR alfa ($p < 0,01$) e beta ($p < 0,001$) no grupo não metastático. Não houve diferença na carga de mutação e neoantígenos entre os grupos. Contudo, a amostra com maior carga de neoantígenos foi também a única das 22 com uma mutação na via de apresentação de antígenos. Ao todo, foram encontradas quatro amostras com mutações em proteínas da via de reparo ao dano de DNA. Estas quatro amostras possuem a maior carga de neoantígeno entre todas as amostras analisadas. Parece haver uma associação dos linfócitos T CD4 de memória com o desfecho livre de metástase. A análise individual de cada amostra indica mecanismos putativos de escape imunológico, como por exemplo, a mutação em uma proteína da via de apresentação de antígeno.

Palavras-chave: Tumor de cavidade oral, metástase, evasão imunológica



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

Metastização em tumores de cavidade oral HPV negativos: uma abordagem *in silico*

Marco Antônio Marques Pretti

Abstract

Oral cavity tumors are the 5th more frequent neoplasms among men according to INCA's last estimative. The occurrence of metastatic processes during tumor development increases the chances of relapse and hamper the treatment. That said, studying the mechanisms that favour metastasis may help identifying biomarkers associated with this process and ultimately improve treatment. Our aim was to analyze *in silico* the HPV-negative tumors from oral cavity comparing those which and without linfonodal metastasis. 20 samples from a collaborator's cohort (AC Camargo - ACC) and 22 samples from The Cancer Genome Atlas (TCGA) were analyzed. The raw data from ACC was filtered and quality checked using FastQC and Trimmomatic, respectively. The ACC reads were then aligned against the human genome GRCh38 using star method and genes were counted with RSEM package. TCGA cohort already possess pre-processed data. Differentially expressed genes were evaluated by DESeq2, the enriched pathways by Webgestalt and tumor microenvironment by xCell. Exclusively for the TCGA available data, the HLA-I alleles were identified by Optitype and subsequent neoantigen prediction was accomplished by netMHCpan. T and B-cell receptors (TCR and BCR) repertoire were identified and analyzed by MiXCR. We identified 186 DEG for the TCGA cohort, 127 for the ACC and 3 DEG shared genes, as which two of them up-regulated in the same condition: PIWIL2 (up-regulated in the non metastatic group) and ADH1B (up-regulated in the metastatic group). The immune population with the highest correlation coefficient was the memory CD4 T-cell (Pearson: -0,78 in ACC and -0,71 in TCGA) which signature is enriched in the non metastatic group. A polymphocyte B signature had an elevated correlation coefficient (Pearson: 0,77 in TCGA cohort) with the metastatic outcome. A higher clone number of TCR alpha ($p < 0,01$) and beta ($p < 0,001$) chains was identified in the non metastatic group. There was no difference between mutation and neoantigen load between groups. Moreover, the sample with higher mutation burden was also the solely bearing a damaging mutation in the antigen processing and presentation pathway. In total, 4 samples had mutations in DNA repair pathways, representing the 4 with higher neoantigen burden. Pathway enrichment analysis as well as correlations between immune populations are currently underway. The results suggest an association between memory CD4 T-cells and metastasis-free disease. The individual analysis of each sample indicates putative mechanisms of immune escape, for example, the mutated protein in the antigen processing and presentation pathway.

Keywords: Oral cavity tumors, metastasis, immune evasion

Lista de Tabelas

Tabela 3.1 Sumário das análises realizadas para cada coorte.	25
Tabela 3.2 Características das amostras da coorte do ACC.	27
Tabela 4.1 Características das amostras da coorte do TCGA.	32

Lista de Figuras

Figura 1.1 Sítios anatômicos dos tumores de cabeça e pescoço.	1
Figura 1.2 Vias de apresentação de antígenos.	6
Figura 1.3 Comparativo da geração de NeoAg por mutações pontuais <i>versus</i> mudanças no quadro de leitura.	11
Figura 1.4 Microambiente tumoral.	14
Figura 1.5 Exemplos de parâmetros genéticos e imunológicos passíveis de análise por meio de metodologias de sequenciamento em larga escala.	22
Figura 3.1 Esquematização do pipeline de análises desenvolvido pelos grupos e utilizado neste trabalho.	26
Figura 4.1 Qualidade das leituras por base do sequenciamento da coorte ACC geradas pelo FastQC.	34
Figura 4.2 Distribuição do tamanho das sequências entre as leituras.	34
Figura 4.3 Análise de componente principal.	35
Figura 4.4 Estimativa de redução da dispersão das contagens normalizadas pelo DESeq2 nas duas coortes.	36
Figura 4.5 Volcano plot dos genes diferencialmente expressos entre a condição com metástase e sem metástase para a coorte do ACC e do TCGA.	37
Figura 4.6 Genes diferencialmente expressos entre as duas coortes.	38
Figura 4.7 Contagem normalizada pelo DESeq2 dos genes <i>PIWIL2</i> , <i>C10orf105</i> e <i>ADH1B</i>	39
Figura 4.8 Mapa de calor dos DEG nas duas coortes.	40
Figura 4.9 Análise de enriquecimento de vias para as duas coortes considerando diferentes bancos de dados.	42
Figura 4.10 Mapa de calor da análise de enriquecimento de população.	44
Figura 4.11 Correlações de Pearson entre as populações do microambiente e entre os grupos metastático e não metastático.	45
Figura 4.12 Repertório de TCR entre os grupos.	46
Figura 4.13 Repertório de BCR entre os grupos.	47
Figura 4.14 Frequência de mutações nas coortes analisadas para cada um dos oito genes mais frequentemente mutados em tumores de cabeça e pescoço.	48

Figura 4.15 Carga de mutação e neoantígeno da coorte do TCGA.	49
Figura 4.16 Perfil mutacional da coorte TCGA.	51
Figura 4.17 Panorâma imunológico da coorte TCGA.	53-54

Lista de Abreviaturas

ACC - AC Camargo Cancer Center
APC - *Antigen-presenting cell*, Célula apresentadora de antígeno
BAM - *Binary Alignment Map* - Mapa de alinhamento binário
BCR - *B-cell receptor*, Receptor de célula B
CAF - *Cancer associated fibroblast*, Fibroblasto associado ao tumor
CDR3 - Complementarity-determining region 3, Região determinante da complementariedade 3
Célula NK - Célula exterminadora natural
CTL - *Cytotoxic T cell*, Linfócito T citotóxico
DAMP - *Damage-associated molecular pattern*, Padrão molecular associado ao dano
DEG - *Differentially expressed genes*, Genes diferencialmente expressos
EBV - Epstein-Barr virus
EMT - *Epithelial-mesenchymal transition*, Transição epitélio-mesênquima
ER - *Endoplasmic reticulum*, Retículo endoplasmático
ERAAP - Aminopeptidase ER 1
FcR - Receptor de Fc
HLA - *Human Leukocyte Antigen*, Antígeno leucocitário humano
HPV - *Human Papillomavirus*, Vírus do papiloma humano.
HSC - *Hematopoietic stem cell*, Célula-tronco hematopoiética
IDO - indolamina 2,3-dioxigenase
Ig - Imunoglobulina
Ii - *Invariant chain*, Cadeia invariante
Indel - Inserção e deleção
KEGG - *Kyoto Encyclopedia of Genes and Genomes*
KIR - *Killer-cell immunoglobulin-like receptors*, Receptor de Ig de célula NK
MDSC - *Myeloid-derived suppressor cell*, Célula supressora de origem mieloide
MMR - *Mismatch repair*, Reparo de pareamento errôneo
NeoAg - Neoantígeno
PAMP - *Pathogen-associated molecular pattern*, Padrão molecular associado à patógeno
Panther - *Protein ANalysis THrough Evolutionary Relationships*
PCA - *Principal component analysis*, Análise de componente principal
p:HLA - Complexo peptídeo:HLA
SI - Sistema imunológico
TAA - *Tumor associated antigens*, Antígenos associado ao tumor
TAP - *Transporter associated with antigen processing*, Transportador associado ao processamento de antígeno
TCGA - *The Cancer Genome Atlas*
TCR - *T-cell receptor*, Receptor de célula T
Th - *Thelper cell*, Linfócito T auxiliar
TLR - *Toll-like receptor*, Receptor do tipo Toll
TPM - Transcritos por milhão
Treg - Linfócito T regulador
TSA - *Tumor Specific Antigens*, Antígenos específicos de tumor
VCF - *Variant Call Format* - Formato de chamada de variantes
VEGF - *Vascular Endothelial Growth Factor*, Fator de crescimento endotelial vascular

Sumário

Resumo	VI
Abstract	VII
Lista de Tabelas	VIII
Lista de Figuras	IX
Lista de Abreviaturas	XI
1 Introdução	1
1.1 Imunologia de tumores	3
1.2 Antígenos tumorais	8
1.2.1 Antígenos associados a tumor	8
1.2.1.1 Antígenos carcinoembrionários e testiculares tumorais	8
1.2.1.2 Antígenos virais	9
1.2.1.3 Outros antígenos associados a tumor	9
1.2.2 Antígenos específicos de tumor - neoantígenos	10
1.3 Eliminação das células tumorais pelo sistema imunológico	12
1.3.1 Morte por Linfócito T CD8 ativado	13
1.3.2 Morte por célula NK	15
1.3.3 Morte por macrófagos	16
1.4 Evasão tumoral	16
1.4.1 Evadindo o reconhecimento imunológico	17
1.4.2 Impedindo a atividade efetora do sistema imunológico	17
1.4.3 Recrutamento de células inibitórias para o microambiente tumoral	18
1.5 Inflamação no tumor: um vilão necessário	19
1.6 Indução da angiogênese e metástase	21
1.7 Imunoinformática: imunologia na era do NGS	22
2 Objetivo	24
2.1 Objetivo geral	24
2.2 Objetivos específicos	24
3 Material e Métodos	25
3.1 Obtenção das amostras	26
3.1.1 Coorte selecionada pelo AC Camargo	26
3.1.2 Critérios de inclusão da coorte do TCGA	28
3.2 Pré-processamento dos dados	28
3.3 Expressão diferencial e enriquecimento de vias	29
3.4 Carga de mutação e predição de ligação dos peptídeos ao HLA	29
3.5 Identificação de TCR e BCR	30
3.6 Estimativa das populações do microambiente tumoral	30
3.7 Análises estatísticas e geração de dados em R	31
4 Resultados	32
4.1 Características das amostras selecionadas	32
4.2 Pré-processamento dos dados	33
4.2.1 Processamento e filtragem dos dados brutos	33
4.2.2 Efeito de lote da coorte ACC	35
4.3 Genes diferencialmente expressos e vias enriquecidas	36
4.4 Populações do microambiente tumoral	43
4.5 Repertório de TCR e BCR	45
4.6 Perfil mutacional e de neoantígenos	48
4.7 Escape tumoral	52
5 Discussão	55
6 Conclusão	69
7 Referências Bibliográficas	71

8 Anexos	87
Anexo 1 - Lista dos genes de reparo ao dano no DNA pesquisados	87
Anexo 2 - Lista dos genes da via de processamento e apresentação pesquisados	88
Anexo 3 - Lâmina do paciente com estrutura sugestiva de órgão linfoide terciário	89
Anexo 4 - Carta de aprovação do Comitê de Ética em Pesquisa do AC Camargo Cancer Center	90

1 Introdução

Câncer é um nome genérico dado a um conjunto de mais de 100 doenças que possuem mecanismos de crescimento e de morte desordenados (NIH, 2015; INCA, 2018b). As principais características das células tumorais foram revisadas por Hanahan e Weinberg e incluem, por exemplo, resistência à apoptose, evasão do sistema imunológico e indução da angiogênese (HANAHAN; WEINBERG, 2000, 2011). O câncer é a segunda doença que mais causa morte no mundo (WHO, 2018), com estimativa de mais de 9,5 milhões de novos casos em 2018 (IARC, 2018), e a previsão é que se torne a primeira causa de morte nos Estados Unidos da América em 2020 (WEIR et al., 2016). A estimativa do Instituto Nacional de Câncer para o Brasil em 2018 é de 582.590 novos casos de câncer (INCA, 2018a). Conforme essa mesma estimativa, o câncer de cavidade oral é o quinto na lista dos mais incidentes entre os homens, excluindo-se câncer de pele não melanoma. Essa alta incidência e mortalidade, o torna relevante para a saúde pública do Brasil (NIH, 2014).

Os tumores da cavidade oral incluem aqueles da região da boca, lábios, céu da boca, bochecha, língua e região sublingual (INCA, 2018c). A cavidade oral é o principal sítio anatômico afetado por tumores de cabeça e pescoço, que também incluem as glândulas salivares, laringe, faringe e a cavidade nasal, ilustradas na Figura 1.1 (NIH, 2017).

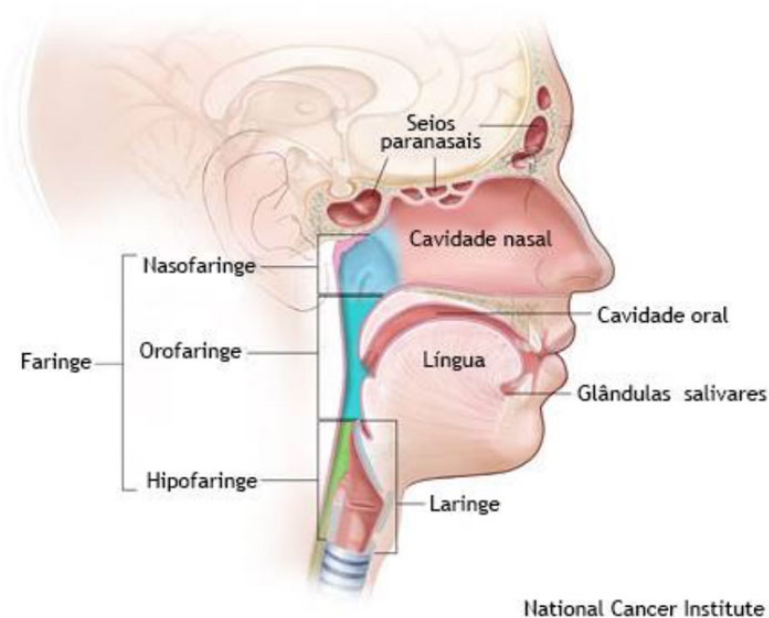


Figura 1.1 Sítios anatômicos dos tumores de cabeça e pescoço (NIH, 2017).

Seus principais fatores de risco são o tabagismo, o consumo de álcool e a infecção por HPV - *Human Papilloma virus*, apesar desta última ser particularmente associada a tumores da orofaringe (AMERICAN CANCER SOCIETY, 2018a). A idade média dos pacientes afetados é de 60 anos para homens e de 67 para mulheres; esta diferença pode ter relação com os fatores de risco supracitados que foram identificados como sendo mais frequentes em homens (PIRES et al., 2013). Os tumores de cavidade oral se manifestam, em mais de 90% dos casos, sob a forma de carcinoma de células escamosas, um crescimento anormal das células epiteliais da camada mais superior do tecido (AMERICAN CANCER SOCIETY, 2018b).

O estágio inicial destes tumores, que acometem a camada mais superior do epitélio, é chamado de carcinoma *in situ*. (AMERICAN CANCER SOCIETY, 2018b). Durante a progressão do tumor, as células tumorais podem exibir fenótipos mais invasivos - como mudanças na interação com a matriz extracelular e maior motilidade - e romper a lâmina própria, invadindo camadas mais interiores do tecido. Neste processo, células tumorais podem ser capturadas por capilares linfáticos e atingir linfonodos cervicais (HEERBOTH et al., 2015). O tamanho tumoral e a presença de metástases linfonodais são dois dos principais fatores que diminuem consideravelmente a sobrevida dos pacientes. Tumores classificados como estágio I e II são geralmente submetidos à cirurgia ou radiação enquanto tumores em estádios mais tardios (III e IV) são removidos cirurgicamente e submetidos a tratamentos adjuvantes (BELCHER et al., 2014). O procedimento realizado em casos de metástase nos linfonodos cervicais é a retirada dos mesmos, procedimento invasivo conhecido como esvaziamento cervical. Ele é responsável por diminuir a progressão da doença e é realizado normalmente para tumores em estádios mais avançados com evidência de metástase por meio de biópsia linfonodal. Contudo, em alguns casos, o clínico pode sugerir cirurgias eletivas de esvaziamento cervical baseando-se no tamanho do tumor, mesmo sem evidências de invasão linfonodal (BELCHER et al., 2014; AMERICAN CANCER SOCIETY, 2018c).

De modo a identificar melhores marcadores da invasividade de tumores da cavidade oral, Netto e colaboradores realizaram um trabalho comparando (i) tumores classificados patologicamente em t1 e t2 com presença de metástase linfonodal *versus* (ii) tumores classificados patologicamente em t3 e t4 sem presença de

metástase linfonodal. No trabalho, foram identificados transcritos com potencial de marcador de metástase, dentre os quais o *NR3C2* foi o mais promissor (NETTO, 2015).

Todavia, não foram exploradas possíveis participações imunológicas e do microambiente que possam ter contribuído com a progressão, evasão e invasividade tumoral. Desde o começo do século passado, hipóteses foram elaboradas sobre a participação do Sistema Imunológico (SI) na eliminação de tumores (EHRlich, 1909; THOMAS; LAWRENCE, 1959; BURNET, 1970). Contudo, somente por volta dos anos 2000, evidências experimentais surgiram para corroborar essas teorias (DUNN et al., 2002). O modelo mais aceito atualmente é composto por três fases: durante a primeira fase o SI elimina continuamente novas células tumorais (fase de eliminação, em geral consequência da imunovigilância); o SI exerce uma pressão seletiva sobre as células tumorais de modo a selecionar as menos imunogênicas (fase de equilíbrio); e, após um período de latência que pode durar anos, essas células podem evadir da vigilância imunológica - fase de escape (DUNN et al., 2002). Outro fator que demonstra a importância do SI no controle de tumores é, por exemplo, a presença de infiltrado imunológico no tumor e sua associação com a sobrevida livre de doença em tumores colorretais (PAGÈS et al., 2010).

1.1 Imunologia de tumores

O SI é um grupo de células, órgãos e moléculas que atuam em conjunto na manutenção da fisiologia tecidual, o que inclui o combate e eliminação de agentes infecciosos (MURPHY, 2012a). Os componentes do SI e o tipo de resposta que produzem são classicamente divididos em dois braços: a imunidade inata, componentes celulares configurados para uma resposta imediata ao insulto, com repertório de reconhecimento limitado; e a imunidade adaptativa, mais tardia e capaz de um reconhecimento específico por parte de seus receptores, já estes possuindo um repertório quase infinito devido à sua geração a partir de rearranjos de segmentos gênicos. Apesar dessa divisão didática, os braços atuam em sinergia e conjunto, sendo difícil dissociar e classificar uma resposta como totalmente inata ou adaptativa (SILVERSTEIN, 2009).

O SI é capaz de distinguir moléculas que pertencem ao organismo (próprias) das que não pertencem (não próprias) por meio de uma variedade de receptores. A

imunidade adaptativa possui dois receptores para o reconhecimento de moléculas não próprias, o receptor de células B (BCR) e o receptor de células T (TCR). Ambos os receptores são gerados a partir de uma recombinação gênica durante a maturação dos linfócitos, conferindo uma diversidade de receptores estimada em 5×10^{13} para BCR e 10^{18} para TCR (MURPHY, 2012b). Moléculas não próprias reconhecidas por esses receptores são denominadas antígenos, termo que, inicialmente, referia-se somente à moléculas que se ligavam com especificidade a anticorpos, mas atualmente inclui o reconhecimento por TCR. Uma proteína viral, por exemplo, é um antígeno e pode ter diferentes porções proteicas reconhecíveis; a cada porção reconhecível dá-se o nome de epítopo (MURPHY, 2012c, 2012a). Alguns dos receptores da imunidade inata são capazes de reconhecer padrões associados a patógenos - PAMP (*Pathogen-associated molecular pattern*, do inglês Padrões moleculares associados a patógenos) e moléculas associadas à quebra de fisiologia celular, as chamadas DAMP (*Damage-associated molecular patterns*, do inglês Padrões moleculares associados ao dano). Um mesmo receptor pode não somente reconhecer um ligante exógeno (PAMP), mas também sinalizar quebra de fisiologia via DAMP. O TLR-4, por exemplo, é um dos muitos receptores de padrão presentes no braço da imunidade inata cuja ativação facilita a fagocitose de patógenos e liberação de citocinas¹; ele pode ser ativado por lipídios da parede bacteriana ou por HMGB1, uma molécula própria (BLANDER; MEDZHITOV, 2006; PICCININI; MIDWOOD, 2010).

Se por um lado os receptores associados à imunidade inata reconhecem padrões pré-determinados, o TCR e BCR são gerados aleatoriamente sem alvo pré-definido. Isso permite, teoricamente, que o TCR e BCR reconheça qualquer epítopo ou antígeno, inclusive epítopos próprios, não fosse a exclusão de receptores autorreativos durante o desenvolvimento linfocitário (tolerância central) e mecanismos de tolerância periférica a antígenos (NEMAZEE, 2006).

O reconhecimento de uma célula tumoral via TCR ocorre por meio do complexo de histocompatibilidade principal, que em humanos recebe o nome de HLA (*Human leukocyte Antigen*). O reconhecimento de um complexo peptídeo:HLA-I (p:HLA) por um TCR de um linfócito T CD8 previamente primado desencadeará mecanismos citotóxicos. Os antígenos das células tumorais - na maioria peptídeos -

¹ moléculas solúveis que modulam a resposta de células do SI

são reconhecidos por meio de receptores celulares (BCR e TCR) ou solúveis (anticorpos) (FREMD et al., 2013; REINHERZ, 2015). O HLA são compostos por duas classes principais: HLA de classe I que interage com TCR de linfócitos do subtipo CD8; e HLA de classe II que interage com TCR de linfócitos T CD4, conforme a Figura 1.2-A (MURPHY, 2012d). As diferenças entre as duas classes de HLA vão além, o HLA-I é composto por uma cadeia alfa e uma molécula de beta-2-microglobulina (β_2M), já o HLA-II é formado por uma cadeia alfa e uma beta. As moléculas de HLA possuem uma fenda em sua superfície destinada à ligação de peptídeos. No HLA de classe I essa cavidade acomoda peptídeos de 7 a 13 resíduos (7 a 13 mer, normalmente 9 mer), contrastando com o HLA-II que acomoda peptídeos de tamanhos maiores, de 11 a 30 mer (SCHUMACHER et al., 1991; RAMMENSEE; FRIEDE; STEVANOVIĆ, 1995). Ao contrário da maioria dos genes, o HLA é poligênico e polimórfico, de modo que existem 3 pares de genes que codificam HLA de classe I (HLA-A, -B e -C) e 3 para classe II (HLA-DP, -DQ e -DR). Um indivíduo pode possuir até 6 alelos diferentes de HLA-I e outros 6 de HLA-II. Essa variedade de moléculas de HLA é importante, pois a estas moléculas se ligarão peptídeos intracelulares (HLA-I) ou extracelulares (HLA-II), próprios ou não próprios, e essa ligação dependerá da afinidade do peptídeo pelo HLA, que difere conforme o alelo (MURPHY, 2012e). Todas as células nucleadas possuem HLA-I, mas somente algumas células expressam HLA-II - e.g. macrófagos, células dendríticas e linfócitos B, células designadas como apresentadoras profissionais de antígeno.

O trajeto de um peptídeo até a fenda do HLA-I tem início com a degradação de proteínas pelo complexo do proteassomo e geração de peptídeos de tamanhos variados; eles serão transportados para o retículo endoplasmático através da proteína TAP (*Transport Associated Protein*, Proteína associada ao transporte) a fim de encontrar a molécula de HLA; em seguida sofrem ação da aminopeptidase ERAAP que clivará as extremidades dos peptídeos de modo a facilitar seu encaixe na fenda; por fim, os peptídeos são carregados no HLA-I com auxílio de um complexo proteico que tem participação da tapasina, calreticulina e β_2M (MURPHY, 2012f) (Fig 1.2-A). Nessa última fase, há uma competição entre os peptídeos presentes no retículo para ligação às moléculas de HLA-I. Já os peptídeos apresentados via HLA-II nas APC tem origem extracelular; eles são endocitados e

degradados em vesículas as quais se fundem com outras vesículas contendo moléculas de HLA-II (Fig 1.2-B).

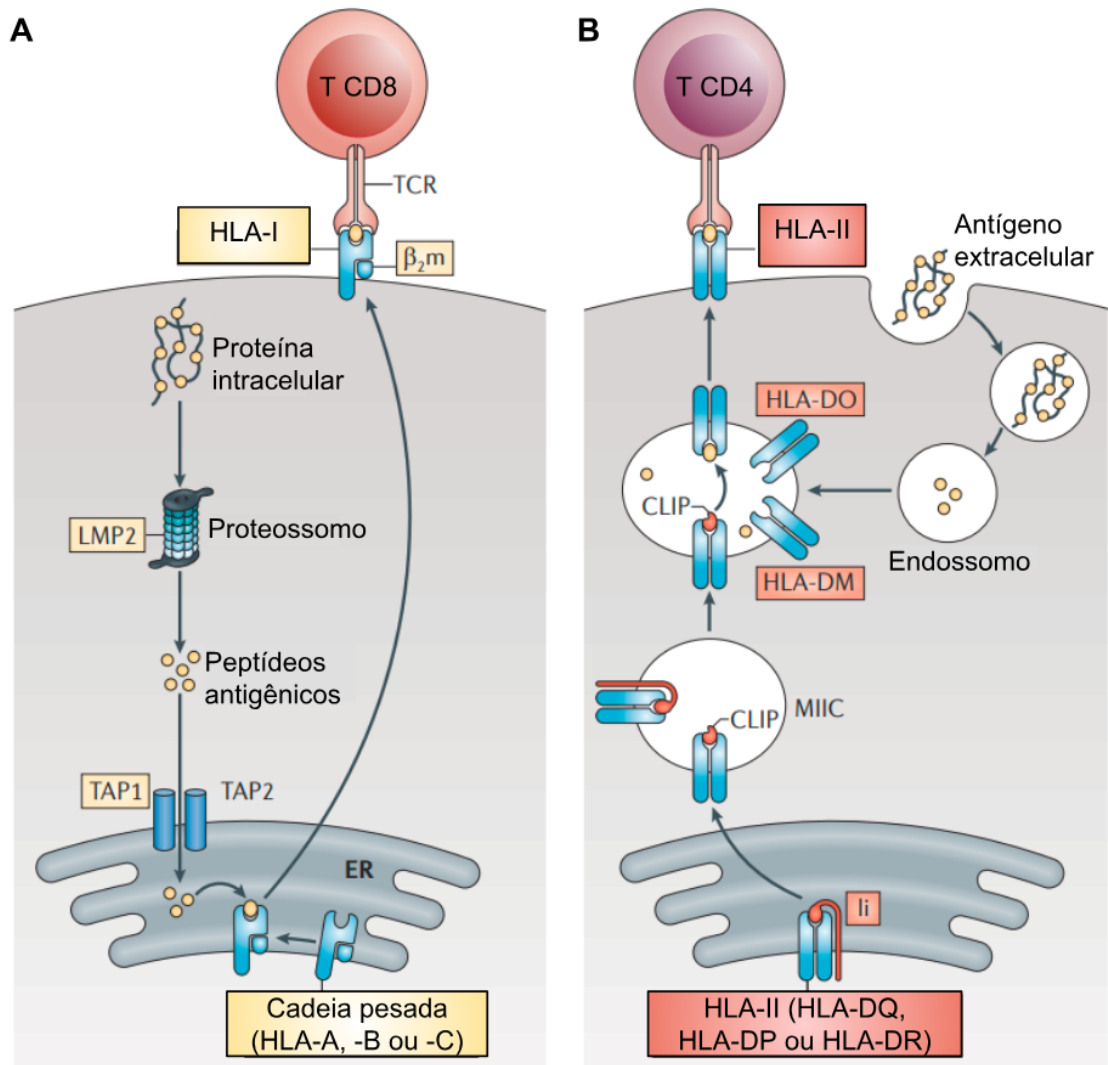


Figura 1.2 Vias de apresentação de antígenos. **A.** O processamento e a apresentação via HLA-I na qual proteínas intracelulares são degradadas pelo complexo do proteossomo (LMP2) em peptídeos menores e então carregados pela TAP (TAP1/TAP2) do citosol ao retículo endoplasmático (ER) onde serão processados pela ERAP em peptídeos menores (não ilustrados) e competem pelo carregamento na fenda do HLA-I; O complexo p:HLA-I é transportado para a membrana onde será testado pelos receptores (TCR) dos linfócitos T CD8. **B.** A apresentação via HLA-II na qual peptídeos extracelulares são endocitados e processados no próprio endossomo. A fusão de dois endossomos ocorre, um contendo os peptídeos extracelulares e o outro contendo as moléculas de HLA-II com a fenda ocupada pela proteína CLIP que logo vai sendo substituída por antígenos com afinidade à fenda. TCR - Receptor de célula T. β_2M - Beta-2-microglobulina. ER - Retículo endoplasmático. MIIC - Compartimento do HLA-II. Ii - Cadeia invariante. Adaptado de (KOBAYASHI; VAN DEN ELSEN, 2012).

Ao contrário da molécula de HLA-I, o HLA-II sai do retículo endoplasmático com uma molécula estabilizadora acoplada na sua fenda, a cadeia invariante (Ii). Após a fusão da vesícula contendo os HLA-II e a vesícula com os peptídeos extracelulares, haverá a troca da porção da cadeia invariante em contato com a fenda (CLIP) por um peptídeo e então a molécula de HLA-II seguirá para a membrana celular. Os peptídeos e antígenos tumorais serão discutidos a fundo em outro capítulo.

A eliminação de uma célula tumoral por um linfócito T CD8 depende da ativação prévia do linfócito, que ocorre normalmente no linfonodo e por intermédio de uma célula apresentadora de antígeno (APC). Uma APC expressa HLA de classe I e II, bem como moléculas co-estimulatórias necessárias à ativação do linfócito. Elas estão presentes nos tecidos e endocitam material extracelular ou mesmo fagocitam células que exibem sinais de estresse ou morte celular (e.g. via DAMP). Proteínas como a HMGB1 e calreticulina, e moléculas como o ATP são alguns exemplos de DAMP liberados para o meio extracelular devido a diversos tipos de morte celular imunogênicos e podem ativar as APC (MELCHER et al., 1999; HERNANDEZ; HUEBENER; SCHWABE, 2016). Quando a APC é ativada por receptores da imunidade inata, a endocitose diminui e a expressão de moléculas co-ativadoras e receptores de quimiocinas aumenta. A APC então migra para o linfonodo, onde linfócitos T *naïve*² testam a interação dos TCR com os p:HLA-I das células dendríticas (GUERMONPREZ et al., 2002). A ativação do linfócito - conhecida como *priming* - depende então do contato de um receptor de linfócito T com um p:HLA ao qual ele é reativo (1º sinal); da interação entre moléculas de superfície do linfócitos e moléculas co-estimulatórias da APC (2º sinal); e do estímulo conferido por citocinas (3º sinal), esse último influenciará no fenótipo do linfócito ativado. Se houver uma interação forte o suficiente para ativar a célula T, a mesma irá sofrer uma expansão clonal no órgão linfoide durante 4 a 6 dias a fim de gerar células efetoras e de memória. Após a expansão, o linfócito deixa o órgão linfoide e retorna à circulação sanguínea, através da qual irá trafegar rumo aos sítios inflamados em busca de seus alvos - no caso dos tumores, o p:HLA-I das células tumorais ao qual é reativo.

Nas células em geral, as moléculas de HLA-I contêm antígenos intracelulares, mas nas APC ocorre um fenômeno conhecido como apresentação cruzada. Nele,

² linfócitos maduros que ainda não encontraram um antígeno ao qual são reativos

antígenos extracelulares captados, que supostamente são apresentados via HLA-II, podem ser carregados em moléculas de HLA-I. Isso permite ao SI eliminar células tumorais e infecções virais por intermédio dos linfócitos T CD8 (BELZ; CARBONE; HEATH, 2002). Contudo, existem evidências de melanomas expressando moléculas de HLA-II. Logo, em alguns casos, os linfócitos T CD4 poderiam ter uma participação na eliminação direta do tumor, além de dar suporte à atividade dos linfócitos T CD8 (DEFFRENNES et al., 2001; JOHNSON et al., 2016).

1.2 Antígenos tumorais

Os antígenos tumorais são antígenos expressos pelas células tumorais capazes de desencadear uma resposta imunológica (MURPHY, 2017). Esses antígenos são principalmente classificados como antígenos associados a tumor (TAA, do inglês *Tumor Associated Antigen*) - compartilhados por diferentes tumores e tecidos saudáveis, mas não exclusivos dos tumores - ou antígenos específicos do tumor (TSA, do inglês *Tumor Specific Antigen*), raramente compartilhados por tumores diferentes e tecidos saudáveis (VIGNERON, 2015).

1.2.1 Antígenos associados a tumor

1.2.1.1 Antígenos carcinoembrionários e testiculares tumorais

Os antígenos carcinoembrionários são principalmente expressos durante o período embrionário, apesar de alguns tecidos adultos possuírem algumas dessas moléculas (THOMPSON; ZIMMERMANN, 1988). Já os antígenos testiculares tumorais tem sua expressão restrita a órgãos imunoprivilegiados tais como as células germinativas masculinas (ZHAO et al., 2014). Durante o processo oncogênico, a hipometilação global de promotores de alguns fatores de transcrição pode acarretar a reexpressão dessas moléculas em tecidos saudáveis adultos (DE SMET et al., 1996; JANG et al., 2001; KANG et al., 2007; ALMSTEDT et al., 2010). A princípio, o SI não desenvolve tolerância central para esses antígenos, ou seja, não há depleção de linfócitos reativos para os mesmos, o que os torna imunogênicos quando expressos fora do contexto embrionário ou em tecidos que não os órgãos imunoprivilegiados. O NY-ESO-1, por exemplo, é um antígeno testicular que, quando expresso em tecidos extra-testiculares, é capaz de elicitar uma resposta mediada por anticorpos - via reconhecimento direto - e de células T - via apresentação de

antígenos por HLA (GNJATIC et al., 2006). Por conta disso, a expressão dessas moléculas em alguns tumores pode ser alvo de terapias baseadas em vacinação (SCANLAN et al., 2002).

Além do caráter imunogênico, os TAA podem exercer um papel chave na tumorigênese. Por exemplo, alguns TAA da família MAGE-A foram relacionados a uma pior resposta às terapias de bloqueio de *checkpoint* imunológico³ em melanomas, pois os mesmos estavam associados a mecanismos de inibição da autofagia (SHUKLA et al., 2018). Outros TAA foram associados à manutenção do fenótipo de célula-tronco tumoral e também à regulação positiva de genes que favorecem a transição epitélio-mesênquima e, conseqüentemente, a metástase (YANG et al., 2015). Isso demonstra que esses antígenos estão efetivamente exercendo um papel na célula tumoral, embora esta contribuição certamente tenha sido amplamente negligenciada até então.

1.2.1.2 Antígenos virais

Algumas infecções virais são fatores de risco para o desenvolvimento de tumores, como o EBV (do inglês, *Epstein-Barr virus*) para linfoma de células B ou o HPV para câncer de orofaringe e cervical (FARIDI et al., 2011; CANCER RESEARCH UK, 2018). Ao infectar células normais, esses vírus podem se integrar no genoma da célula causando instabilidade genética além de inibir proteínas repressoras tumorais como p53 e pRb (FARIDI et al., 2011). Por conta do caráter não-próprio das proteínas virais, elas têm potencial para desencadear uma resposta imunológica. Terapias de vacinação que utilizam peptídeos virais tem se mostrado eficazes para a prevenção do câncer (NATIONAL CANCER INSTITUTE, 2018).

1.2.1.3 Outros antígenos associados a tumor

Uma outra classe de antígenos são os antígenos de diferenciação, pois sua expressão é restrita a um tecido ou tipo celular. As proteínas tirosinase, MART1 e gp100 estão superexpressas em melanomas e são específicas da linhagem melanocítica (MURPHY, 2017). Em condições normais sua expressão não é suficiente para induzir uma resposta imunológica, mas com a superexpressão, a relação de p:HLA-I contendo peptídeos derivados dessas proteínas aumenta,

³ terapia que utiliza anticorpos para bloquear moléculas inibitórias da resposta imunológica.

possivelmente supera o limiar de ativação, e é capaz de elicitar respostas de linfócitos T CD8 e CD4 (MURPHY, 2017). O fato de um antígeno próprio estar expresso muito acima dos níveis de expressão de uma célula normal também pode desencadear uma resposta imunológica contra ele pelo mesmo mecanismo dos antígenos de diferenciação. Esse é o caso do HER-2/neu em tumores de mama, ovário, colorretais dentre outros (CORREA; PLUNKETT, 2001). Outros antígenos próprios se tornam imunogênicos devido à modificações pós-transcricionais diferentes, como é o caso da proteína MUC1 que se torna menos glicosilada em tumores (NATH; MUKHERJEE, 2014).

1.2.2 Antígenos específicos de tumor - neoantígenos

Um subtipo particular de TSA são os neoantígenos (NeoAg), epítomos de célula T gerados a partir de mutações somáticas (SCHUMACHER; SCHEPER; KVISTBORG, 2018). As mutações somáticas podem ocasionar a formação de novos epítomos, os chamados NeoEpítomos, de acordo com o impacto da variante na proteína. Conforme o prefixo “neo” indica, eles são novos epítomos que podem ou não ser reconhecidos pelos receptores do SI. Contudo, nem todo NeoEpítomo irá desencadear uma resposta imunológica, pois alguns pré-requisitos são necessários para tal. Primeiramente, é preciso que a proteína que gera o peptídeo mutado esteja expressa para que o NeoEpítomo seja produzido (WIRTH; KÜHNEL, 2017). Em segundo lugar, é necessária a ligação do peptídeo mutado à fenda das moléculas de HLA com uma afinidade mínima para que ali ele permaneça estável. Por fim, uma vez na membrana celular, o complexo p:HLA-I contendo o NeoEpítomo deverá ser imunogênico e ativar o TCR de um linfócito (WIRTH; KÜHNEL, 2017). Neste texto, NeoEpítomo e NeoAg serão tratados como sinônimos.

Os NeoAg são gerados a partir de mutações somáticas em regiões codificantes que geram a mudança de, pelo menos, um aminoácido na proteína a qual codificam (HEEMSKERK; KVISTBORG; SCHUMACHER, 2013). Mutações sinônimas, que ocasionam troca de códon mas não do aminoácido, não geram NeoAg. Variantes passíveis de gerar NeoAg são os polimorfismos de base única não sinônimos, as inserções e deleções (indel) com ou sem mudanças de fase de leitura. A proteína mutada será clivada em peptídeos menores e o peptídeo que contém pelo menos um aminoácido diferente daquele da sequência original tem chance de

ser imunogênico (RAJASAGI et al., 2014). Em alguns contextos, translocações cromossômicas podem gerar NeoAg derivados dessas proteínas de fusão, como é o caso da proteína de fusão BCR-ABL⁴ (CLARK et al., 2001). A grande maioria dos NeoAg é derivada de mutações passageiras - não relacionadas à carcinogênese - e são principalmente geradas em tumores com altas cargas de mutação, como muitos dos tumores do tipo melanoma, de pulmão e esôfago (SCHUMACHER; SCHREIBER, 2015). A alta carga de mutação se mostrou inclusive um parâmetro de prognóstico importante, pois favorece a resposta das terapias de bloqueio de *checkpoint* imunológico (RIZVI et al., 2015).

Os NeoAg, apesar de não-próprios, podem guardar muita similaridade com peptídeos próprios, especialmente se diferem em apenas um aminoácido (Fig 1.3A). Por conta disso, NeoAg derivados de indel e de mudanças na fase de leitura geram NeoAg (Fig 1.3B) com uma afinidade pelo HLA até nove vezes maior que aqueles derivados de mutações pontuais (TURAJLIC et al., 2017). Além disso, as mutações pontuais tendem a gerar menos NeoAg que as mudanças de quadro de leitura e indel.

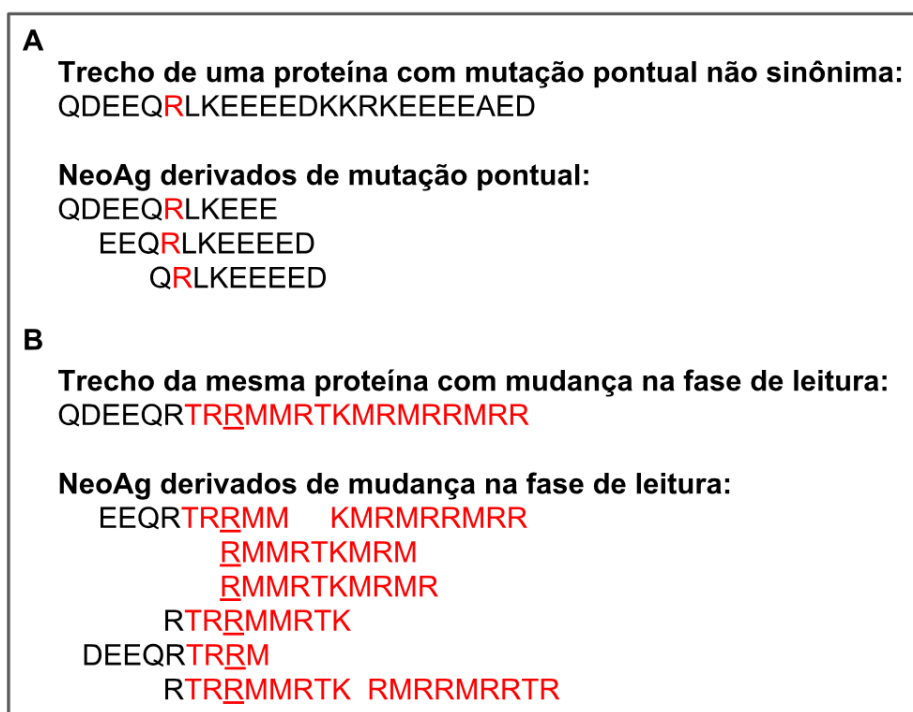


Figura 1.3 Comparativo da geração de NeoAg por mutações pontuais **(A)** versus mudanças no quadro de leitura **(B)** no éxon 9 da proteína CALR. A saída de fase de leitura gera uma sequência amino-terminal nova e com muitas possibilidades para geração de NeoAg. Predição *in silico* considerando os alelos A*02:05, A*03:01, B*18:01, B*35:01, C*05:01 e C*02:02.

⁴ proteína gerada pela fusão do gene *BCR* (cromossomo 22) e *ABL1* (cromossomo 9) na leucemia mieloide crônica

Alguns outros fatores influenciam a produção de NeoAg por aumentarem drasticamente a carga de mutação. Mutações em genes de reparo de DNA, principalmente aqueles da via de reparo de pareamento incorreto (MMR, do inglês *mismatch repair*) ocasionam um fenômeno de hipermutação somática em alguns tumores e, conseqüentemente, elevada carga de NeoAg (GERMANO et al., 2017). Mutações na via de MMR são outro parâmetro associado ao prognóstico favorável da resposta a bloqueio de *checkpoint* imunológico (LE et al., 2017). Foram observadas diversas mutações em enzimas da via de MMR e no gene *POLE* (subunidade da DNA polimerase) em tumores colorretais hipermutados e algumas das enzimas da via de MMR sofriam silenciamento por mecanismos epigenéticos - e.g. MLH1 (CANCER GENOME ATLAS NETWORK, 2012). Proteínas como a ATR⁵, quando mutadas, também já foram implicadas no aumento da carga de NeoAg em melanoma. A ATR é especialmente importante na indução do arresto celular via parceiros como CHK1 e p53 (BRADBURY; JACKSON, 2003; CHEN et al., 2017).

Em adenocarcinoma pulmonar, além de aumentar a quantidade de NeoAg, essas mutações também aumentaram o número de linfócitos infiltrantes no tumor (CHAE et al., 2018). Liepe e colaboradores (LIEPE et al., 2016) demonstraram que uma fração dos peptídeos gerados pelo complexo do proteassomo durante a clivagem proteica é derivada do *splicing* de peptídeos. Isso significa que uma proteína originadora de NeoAg pode ter parte do seu peptídeo mutado combinado com uma região distante da mesma proteína, o que gera peptídeos diferentes que tanto podem ser imunogênicos como não (LIEPE et al., 2016; VIGNERON et al., 2017). Por fim, foi demonstrado que edições de mRNA pela família ADAR (trocas de Adenosina para Iosina) são capazes de induzir resposta de linfócitos T CD8 específicas para o NeoAg editado (ZHANG et al., 2018).

1.3 Eliminação das células tumorais pelo sistema imunológico

Como dito, a eliminação de uma célula tumoral depende primeiramente de seu reconhecimento pelo SI como não-própria, fenômeno que ocorre geralmente por meio da interação do TCR de um linfócito T CD8 previamente ativado com o p:HLA-I da célula tumoral carregado com um antígeno tumoral (MORRIS et al., 2005). A morte de uma célula tumoral pode ser fruto do reconhecimento por uma célula NK,

⁵ proteína sensora de danos à fita de DNA ocasionados por radiação UV

ocasionado pela quimioterapia ou mesmo devido aos danos ao DNA e estresse oxidativo aos quais as células tumorais estão intrinsecamente submetidas (GASSER et al., 2005; PROSKURYAKOV; GABAI, 2010; AOTO et al., 2018). Esses fatores colaboram com a morte imunogênica e favorecem a atividade antitumoral do SI.

1.3.1 Morte por Linfócito T CD8 ativado

O linfócito T CD8 ativado recebe o nome de CTL (do inglês, *Cytotoxic T Cell*), célula com propriedades citotóxicas. Infiltrados imunológicos em tumores ricos em CTL estão relacionados a um melhor prognóstico (PAGÈS et al., 2010; YU et al., 2018). Por outro lado, o linfócito T CD4 ativado pode assumir fenótipos variados conforme o estímulo que recebeu durante a ativação: T *helper* 1 - Th1, Th2, Th17, etc. O perfil Th1 suporta a atividade citotóxica do CTL enquanto o perfil Th2 favorece a ativação de linfócitos B (HOSOI et al., 2008; HADRUP; DONIA; THOR STRATEN, 2013). Ainda que a maioria dos tumores sólidos não possua HLA-II para serem reconhecidos pelos linfócitos T CD4, eles são de grande importância na resposta a tumores hematológicos nos quais o HLA-II é expresso (DADMARZ et al., 1996; MENDEZ et al., 2009). O linfócito T CD4 pode estimular a atividade do CTL diretamente ou por meio de citocinas como o IFN- γ à medida em que reconhece antígenos restritos ao HLA-II (MORRIS et al., 2005; HADRUP; DONIA; THOR STRATEN, 2013). Quando um CTL encontra no tumor um p:HLA-I ao qual é reativo ele é capaz de liberar grânulos contendo perforina e granzima, além de expôr ligantes de morte como FasL e TRAIL (Fig 1.4). A perforina facilita a entrada da granzima na célula tumoral, onde ela cliva seus alvos e induz mecanismos de morte celular; os ligantes FasL e TRAIL se ligam a receptores de morte na superfície da célula tumoral (MARTÍNEZ-LOSTAO; ANEL; PARDO, 2015). A morte por CTL pode expôr mais antígenos da célula tumoral que será eventualmente fagocitada por macrófagos e amplificar a resposta imunológica em curso, seja por re-estímulo com o mesmo antígeno ou novos antígenos. A manutenção de uma resposta antitumoral efetiva é favorecida com essa morte imunogênica - aquela que libera HMGB1 e expõe calreticulina - pois estimula continuamente os linfócitos (ZITVOGEL et al., 2010; AOTO et al., 2018).

Microambiente Tumoral

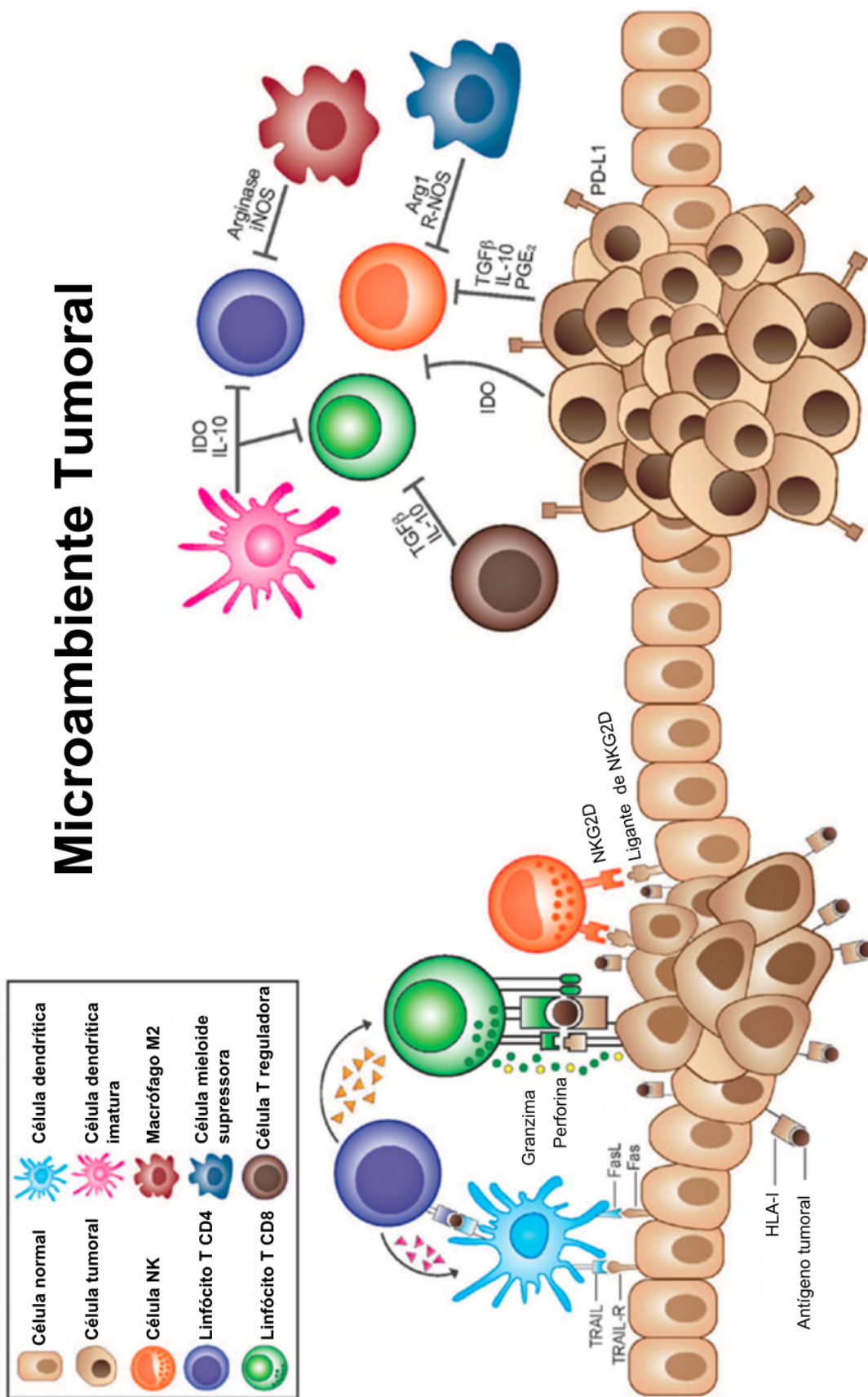


Figura 1.4 Microambiente tumoral mostrando a eliminação de células tumorais por intermédio de CTL (Linfócito T CD8), linfócito T CD4, célula NK e DC (esquerda) e os mecanismos de inibição da atividade dessas células pelo recrutamento de células inibitórias, citocinas e expressão de moléculas de superfície no tumor (direita). Adaptado de: (MONJAZEB et al., 2013).

1.3.2 Morte por célula NK

Componentes da imunidade inata também são capazes de reconhecer e eliminar células tumorais. A célula NK (*natural killer*) possui receptores de membrana de ativação capazes de interagir com ligantes na superfície da célula tumoral. Um deles é o receptor NKG2D que possui diversos ligantes, alguns se tornam superexpressos na célula tumoral em condições de estresse replicativo e quebra da dupla fita de DNA, características mais associadas à células em constante proliferação (GASSER et al., 2005). Outra característica das células tumorais é a regulação positiva de vias como as de Ras, PI3K e MAPK. Uma vez constantemente super-reguladas, induzem a expressão de ULBP1-3, um ligante de NKG2D (LIU et al., 2012). Foi demonstrado que, em um contexto de senescência de células tumorais, o aumento de p53 induziu a expressão de ligantes de NKG2D e de CCL2 - quimiocina que recruta células NK. Esse perfil favoreceu a morte de células tumorais mediada por células NK, conforme ilustrado na Figura 1.4 (IANNELLO et al., 2013).

Além dos receptores ativadores, as células NK também possuem receptores inibitórios, como o KIR (*Killer-cell Ig-like Receptor*). Esse receptor interage com moléculas de HLA-A, -B, -C e -G independentemente do peptídeo apresentado impedindo a ativação da célula NK, mecanismo que ficou conhecido como *missing-self* - ausência do próprio. Esse mecanismo busca eliminar células com regulação negativa de moléculas de HLA-I, o que eventualmente favorece o escape imunológico (LJUNGGREN; KÄRRE, 1990). Outro receptor com função similar é o C94/NKG2 que reconhece um peptídeo sinal apresentado via HLA-E⁶. Esse peptídeo sinal consiste de fragmentos de moléculas de HLA-I e, apresentado dessa forma, indica à célula NK que a célula está expressando as moléculas de HLA-I corretamente (BRAUD et al., 1998). Cada célula NK possui diferentes combinações dos receptores KIR e C94/NKG2 de modo que a ausência de somente um alelo de HLA-I pode ser percebida e ativar os mecanismos de morte da célula NK (RAULET; VANCE; MCMAHON, 2001). As células NK também são responsáveis pelo fenômeno de citotoxicidade celular mediada por anticorpo. O receptor da célula NK FcγR reconhece a porção Fc de anticorpos ligados a seus alvos na superfície de células, o que desencadeará a liberação de grânulos citotóxicos por parte a célula

⁶ molécula não clássica de HLA-I, assim como o HLA-F e HLA-G

NK (WEISKOPF; WEISSMAN, 2015). Após reconhecimento de seus alvos e ativação, a célula NK irá liberar grânulos citotóxicos que contêm principalmente perforina e granzima (LIEBERMAN, 2003; LETTAU et al., 2007).

1.3.3 Morte por macrófagos

Os macrófagos também participam do controle e eliminação de células tumorais, pois possuem receptores que reconhecem padrões associados à morte celular e reconhecimento pela imunidade adaptativa, que podem estar associados à células tumorais. Por exemplo, o receptor para a porção Fc do anticorpo (FcR) na superfície do macrófago reconhece anticorpos complexados a seus alvos na superfície de células, o que induz a fagocitose destas células (WEISKOPF; WEISSMAN, 2015). Outro mecanismo que favorece a fagocitose de células tumorais é a presença na membrana de outros sinais ativadores, como a calreticulina - proteína que controla a homeostase do cálcio intracelular - e a fosfatidilserina, lipídio de membrana da camada intracelular e que funciona como sinal da apoptose quando exposto na membrana extracelular (GARDAI et al., 2005; MIYANISHI et al., 2007). Os macrófagos M1 são grandes produtores de IFN- γ , IL-12 e TNF- α . A IL-12 por exemplo, estimula a diferenciação do linfócito T CD4 no subtipo Th1 que estimulará a atividade fagocítica dos macrófagos (BENOIT; DESNUES; MEGE, 2008). Existe uma associação entre os macrófagos M1 e a atividade antiangiogênica devido ao seu perfil pró-inflamatório e também, possivelmente, devido à ligação a C5a - proteína do sistema complemento (LANGER et al., 2010). Essas citocinas são importantes para estimular a atividade fagocítica do macrófago, mas o mesmo também é capaz de eliminar células tumorais por meio da liberação de óxido nítrico, que vai gerar espécies reativas de oxigênio na célula alvo (KELLER et al., 1990).

1.4 Evasão tumoral

Apesar do arsenal de células e moléculas que se prestam a combater as células tumorais, os tumores conseguem, por vezes, escapar do SI e da morte celular utilizando diversos mecanismos. Alguns destes são intrínsecos da própria célula tumoral, outros envolvem células do microambiente - e.g. fibroblastos, linfócitos - e até mesmo recrutam células imunológicas para o microambiente tumoral. Esse microambiente é composto por células naturais do tecido de origem,

células malignas derivadas do tecido ou de um tecido distante, vasos sanguíneos e células imunológicas residentes ou que infiltraram no tecido após um processo inflamatório (MONJAZEB et al., 2013).

1.4.1 Evadindo o reconhecimento imunológico

Dentre os mecanismos que favorecem a evasão está a regulação negativa da molécula de HLA-I que já foi observada em diversos tumores. A menor apresentação de antígenos via HLA-I dificulta o reconhecimento de antígenos tumorais pelos TCR, mas pode favorecer a morte por células NK pelo mecanismo de *missing-self* (HICKLIN; MARINCOLA; FERRONE, 1999). Moléculas da via de processamento de antígeno como a TAP1/2 e proteínas do complexo proteassomo foram detectadas em níveis muito baixos em células tumorais e já foi demonstrado que a deficiência na proteína TAP1/2 contribui para o escape tumoral (RESTIFO et al., 1993; JOHNSEN et al., 1999). Estudos observaram a perda homozigótica da β_2M , proteína que estabiliza o HLA-I na membrana celular, em melanoma metastático e câncer de pulmão além de mutações que geraram um códon de parada precoce (DEL CAMPO et al., 2014; GETTINGER et al., 2017). Esses mecanismos de evasão dificultam ou impedem o reconhecimento dos antígenos tumorais e posterior eliminação da célula tumoral pelos CTL (MAEURER et al., 1996). Evidências recentes de heterogeneidade tumoral revelaram que o tumor é composto por células com mutações exclusivas e outras compartilhadas (BURRELL et al., 2013). Isso permite que células mais imunogênicas sejam selecionadas negativamente pelo SI - imunoedição - e que o crescimento de células menos imunogênicas ganhe espaço na população do tumor. A resposta imunológica contra um antígeno presente em somente parte das células tumorais pode não eliminar completamente o tumor e levar à perda desse antígeno (KHONG; WANG; ROSENBERG, 2004).

1.4.2 Impedindo a atividade efetora do sistema imunológico

Não só o reconhecimento de uma célula tumoral pode ser afetado por mecanismos de evasão, mas também as atividades efetoras após o reconhecimento do tumor. Diferentes moléculas inibitórias são expressas na superfície de tumores e inibem a atividade dos linfócitos T, tais como o CTLA-4, PD-L1 e BTLA. As moléculas inibitórias são importantes no balanço negativo das respostas imunológicas e são

normalmente expressas nos linfócitos e em poucos outros tecidos (PEGGS; QUEZADA; ALLISON, 2009). O CTLA-4, por exemplo, é expresso em diversas linhagens celulares tumorais e em alguns tumores como melanoma e pulmão (WEBER, 2008; PAULSEN et al., 2017). Tumores de pulmão, ovário, cólon e melanomas podem expressar PD-L1 e evadir da vigilância imunológica (DONG et al., 2002). Níveis elevados dessas e outras moléculas inibitórias estão associados a um pior prognóstico e resposta a terapias de bloqueio a essas moléculas em alguns tumores (HU et al., 2017; XIANG et al., 2018). A interação entre um linfócito *naïve* e a APC em um contexto com níveis insuficientes de moléculas co-estimulatórias e altos de moléculas inibitórias pode gerar um estado de não responsividade - anergia - do linfócito ao antígeno durante a ativação do linfócito no órgão linfóide secundário. APC de tumores e o microambiente tumoral tendem a expressar maiores níveis de moléculas inibitórias e seus ligantes, o que favorece a indução de estados de não responsividade em linfócitos (CRESPO et al., 2013). Outro fenótipo que essas moléculas inibitórias podem induzir é a exaustão dos linfócitos. Linfócitos infiltrados e reativos a antígenos tumorais podem se tornar exaustos quando submetidos a repetidos estímulos pelo antígeno e tendem a regular positivamente a expressão de moléculas inibitórias no linfócito, principalmente PD-1. A persistência do tumor e os níveis crescentes de moléculas inibitórias atenuam a resposta do linfócito e o tornam incapaz de uma resposta efetora (BARBER et al., 2006).

1.4.3 Recrutamento de células inibitórias para o microambiente tumoral

As células tumorais são capazes de recrutar células do SI para o microambiente tumoral a fim de inibir respostas antitumorais em curso através de diferentes citocinas (ZOU et al., 2001; CHAUDHARY; ELKORD, 2016; KUMAR et al., 2016). Os linfócitos T reguladores (Treg), as células mieloides supressoras (MDSC), DC imaturas e os macrófagos estão entre essas células. Os mecanismos inibitórios dessas células divergem e podem ocorrer via citocinas, enzimas ou mesmo por contato direto, como mostrado na Figura 1.4 (MONJAZEB et al., 2013). As principais citocinas inibitórias produzidas pela Treg são o TGF- β e a IL-10, ambas inibem a atividade efetora dos CTL e linfócitos T CD4; o TGF- β pode ainda induzir a geração de Treg induzidas. As Treg consomem grandes quantidades de IL-2, citocina essencial para a sobrevivência dos linfócitos e podem também diminuir a capacidade

das DC de ativar linfócitos (SHEVACH, 2009). A quantidade de Treg infiltrada no tumor está correlacionada com pior sobrevida em diversos tumores, mas aparentemente pode ter um papel controverso em outros (SHANG et al., 2015). As DC imaturas produzem IL-10 e a indolamina 2,3- deoxigenase (IDO), enzima que metaboliza o triptofano em metabólitos imunossupressores (MUNN; MELLOR, 2007). Os macrófagos possuem classicamente dois fenótipos: o fenótipo M1 possui um perfil pró-inflamatório e fagocítico com alta produção de espécies reativas de oxigênio e nitrogênio; e o fenótipo M2 que possui características anti-inflamatórias e pró-angiogênicas, auxiliando no reparo tecidual (WEAGEL et al., 2015). A polarização para o fenótipo M2 pode ser induzido por TGF- β e IL-10, citocinas abundantes no contexto pró-tumoral (MOSSER; EDWARDS, 2008; GONG et al., 2012). Células tumorais também são capazes de produzir TGF- β , IL-10, IDO e prostaglandina E2 a fim de induzir esse microambiente tumoral supressor (MONJAZEB et al., 2013).

1.5 Inflamação no tumor: um vilão necessário

O processo inflamatório ocorre em consequência a um dano tecidual e visa eliminar qualquer agente estranho, que eventualmente tenha penetrado no organismo por conta de uma injúria, e reparar o tecido lesado (COUSSENS; WERB, 2002). Células tumorais podem surgir em decorrência de um processo inflamatório crônico devido à exposição contínua a agentes infecciosos ou carcinógenos (MULTHOFF; MOLLS; RADONS, 2011), mas o tumor também é capaz de iniciar a inflamação. O crescimento exacerbado de células pode danificar e inflamar o tecido, ativando padrões moleculares associados ao dano; tal crescimento pode resultar em hipóxia local e induzir a produção de citocinas que recrutam células inflamatórias, tais como neutrófilos e macrófagos (MUNN, 2017). Diversas citocinas inflamatórias estão associadas a uma melhor ou pior resposta antitumoral e o balanço entre elas é uma das chaves para a eliminação efetiva do tumor. As principais delas são as citocinas pró-inflamatórias TNF- α , IL-6 e IL-8; e as anti-inflamatórias IL-10 e TGF- β (LANDSKRON et al., 2014; SAHIBZADA et al., 2017).

Uma das citocinas mais ambíguas é o TNF- α cuja presença em tumores pode tanto se associar à promoção tumoral como à atividade antitumoral, nesse último caso principalmente quando combinado com IFN- γ ou quimioterapia (BALKWILL,

2002). O TNF- α colaborou com a atividade antitumoral em sarcoma murino quando presente em altas doses (HAVELL; FIERS; NORTH, 1988), mas baixas concentrações por períodos prolongados contribuem para a carcinogênese (BALKWILL, 2006). O local de produção dessa citocina também se mostrou relevante: quando presente no tumor ela exerceu atividades antitumorais, em contraste com sua presença em células estromais, que contribuiu para a progressão tumoral (BALKWILL, 2006). Em tumores de cavidade oral, a exposição prolongada ao TNF- α favoreceu o desenvolvimento de um fenótipo de célula-tronco - que possui maior resistência à terapia (LEE et al., 2012). A concentração da mesma foi superior quando comparou-se tecidos saudáveis com tecidos tumorais e tumores pequenos (t1 e t2) com tumores maiores (t3 e t4), apesar de não significativo neste último caso (NAKANO et al., 1999). O mesmo estudo verificou que os níveis de IL-6 eram maiores em tumores t3 e t4 de cavidade oral e com níveis crescentes quando comparados a tecido saudável, tumor e metástase linfonodal, respectivamente (NAKANO et al., 1999). Assim como o TNF- α , a IL-6 também favoreceu o fenótipo de célula-tronco em tumores (KIM et al., 2013). Ela possui capacidade de promoção da tumorigênese e inibição da apoptose através da sinalização de seu receptor, IL-6R α , que ativa a via JAK/STAT (HODGE; HURT; FARRAR, 2005). A IL-6, juntamente com a IL-8 e TNF- α , apresentam níveis aumentados em tumores de cavidade oral e são inclusive testados como biomarcadores para diagnóstico (SAHIBZADA et al., 2017). A IL-8 está associada com o recrutamento dos neutrófilos para o microambiente tumoral e atividade dos mesmos (WAUGH; WILSON, 2008).

O TGF- β foi descrito em tópicos anteriores pelas suas propriedades imunossupressoras e de indução de Treg, mas em um estadiamento inicial do tumor ele se mostrou importante para a inibição da carcinogênese, tanto *in vitro* como *in vivo* (LANDSKRON et al., 2014). Em um estadiamento mais tardio, o TGF- β promove a transição epitélio-mesênquima, mudança que favorece a metastização (MORRISON; PARVANI; SCHIEMANN, 2013). Nos tumores de cavidade oral o TGF- β 1 elevado teve correlação com a metástase linfonodal, maior resistência à morte celular e ativação da via de IL-6 (CHEN et al., 2012). A IL-10 é outra citocina anti-inflamatória produzida por diversas células do SI e atua inibindo a produção de citocinas pró-inflamatórias, como a IL-6 e TNF- α (OFT, 2014). Além disso, ela diminui a expressão de moléculas co-estimulatórias e de HLA-II nas APC (XI et al., 2017).

Seus efeitos sob os linfócitos T CD8 - principais responsáveis pela eliminação direta de células tumorais - são contraditórios, existem evidências que suportam a inibição desses linfócitos pela IL-10 e outras que indicam o favorecimento da atividade citotóxica (OFT, 2014; XI et al., 2017). Em tumores de cavidade oral, os níveis de IL-10 tiveram correlação com a sobrevida e recidiva nos tumores HPV-positivos, mas não nos HPV-negativos (CHUANG et al., 2012). Tendo isso em vista, o balanço de citocinas pró e anti-inflamatórias no microambiente tumoral e a intensidade do processo inflamatório são cruciais para uma resposta antitumoral efetiva, principalmente em relação às citocinas e quimiocinas com ação angiogênica (COUSSENS; WERB, 2002).

1.6 Indução da angiogênese e metástase

Alguns tumores possuem alta capacidade de ganhar a circulação sanguínea e estabelecer um tumor em um sítio distante, fenômeno conhecido como metástase (NATIONAL CANCER INSTITUTE, 2017). Células não transformadas no tecido de origem do tumor, como fibroblastos, podem contribuir para essa progressão e evasão tumoral a partir da secreção de fatores crescimento, citocinas e da degradação da matriz extracelular (SHIGA et al., 2015). O perfil pró-tumoral dos fibroblastos (CAF) é induzido pelo tumor através da secreção de diversas moléculas, e.g. TNF- α , IL-1 β (NAGASAKI et al., 2014), ou mesmo devido ao estresse mecânico gerado pelo crescimento exacerbado de células no local (RÄSÄNEN; VAHERI, 2010). Uma vez ativados pelo tumor, os CAF produzem altos níveis de IL-6 - citocina com capacidades pró-angiogênicas; e metaloproteases que degradam a matriz extracelular, o que facilita o crescimento tumoral, angiogênese e, eventualmente, a metástase (MUELLER; FUSENIG, 2004; NAGASAKI et al., 2014). Outro fator importante para estimular a angiogênese é o VEGF (fator de crescimento endotelial vascular, do inglês *vascular endothelial growth factor*) amplamente produzido pelos CAF em resposta aos altos níveis de IL-6 e de outros fatores (FGF, PDGF) após a interação com células tumorais (GOMES et al., 2013; SHIGA et al., 2015). Esse conjunto de fatores e citocinas são alguns dos principais componentes que promovem a angiogênese e migração das células tumorais para outro sítios, mas muitas outras moléculas produzidas pelos CAF estimulam o crescimento tumoral sem promover a angiogênese (TAO et al., 2017). A transição epitélio-mesênquima

(EMT), na qual as células vão gradativamente perdendo a morfologia epitelial e ganhando características mesenquimais, faz parte do desenvolvimento normal do organismo, mas também está presente em processos de metástase. Durante essa transição, as células tumorais perdem a adesão celular e sofrem mudanças no citoesqueleto (maior expressão de vimentina e N-caderina) que colaboram com a migração celular (HEERBOTH et al., 2015). Existem evidências de EMT em modelos *in vitro* de tumor de cavidade oral, mas o papel da EMT na metastização parece controverso (ZIDAR et al., 2018).

1.7 Imunoinformática: imunologia na era do NGS

O surgimento do sequenciamento de nova geração (NGS) nos anos 2.000 proporcionou um crescimento exponencial da geração de dados biológicos, principalmente na área oncológica (GAUTHIER et al., 2018; KASAIAN; LI; JONES, 2014). A partir dos dados de NGS foi possível aprofundar o estudo da imunologia de tumores (Fig 1.5) utilizando-se de algoritmos para predição de NeoAg, deconvolução para identificação da proporção de infiltrado imunológico e outras células estromais no tumor, montagem de TCR e BCR, dentre outros (HACKL et al., 2016). A este novo campo da bioinformática deu-se o nome de imunoinformática.

	RNA-Seq	WES	WGS
Análise genômica			
Mutações intrônicas	Verde claro	Verde claro	Verde escuro
Mutações exônicas	Verde claro	Verde escuro	Verde escuro
Variação no número de cópias	Branco	Verde escuro	Verde escuro
Clonalidade tumoral	Verde claro	Verde escuro	Verde escuro
Análise imunológica			
Tipagem do HLA	Verde escuro	Verde escuro	Verde escuro
Predição de neoantígenos	Verde escuro	Verde claro	Verde claro
Expressão gênica	Verde escuro	Branco	Branco
Populações imunológicas	Verde escuro	Branco	Branco
Identificação de receptor de célula T e B	Verde escuro	Verde escuro	Verde escuro

Figura 1.5 Exemplos de parâmetros genéticos e imunológicos passíveis de análise por meio de metodologias de sequenciamento em larga escala. As cores representam a fidelidade e viabilidade das análises para cada fonte de dado (branco -inviável; verde claro - factível, mas não ideal; verde escuro - ideal). WES - *Whole Exome Sequencing*, Sequenciamento de exoma total. WGS - *Whole Genome Sequencing*, Sequenciamento de genoma total. HLA - Antígeno Leucocitário Humano. Referência: (HACKL et al., 2016)

A exploração e análise de tumores alcançou um novo patamar com as análises de imunoinformática. Elas permitem a identificação e pré-seleção de NeoAg mais suscetíveis de serem apresentados pelo tumor, a descrição do tipo de resposta imunológica em curso, a descoberta de marcadores preditivos de resposta a terapias (SHUKLA et al., 2018; XIANG et al., 2018), dentre muitos outros. Graças a essas novas abordagens, o entendimento da biologia tumoral está sendo elucidado em larga-escala. Contudo, não se deve ignorar a participação da experimentação. As análises geradas por bioinformática são mais valiosas quando corroboradas por dados experimentais, uma vez que a ciência experimental e computacional não são excludentes, mas sim, complementares (GAUTHIER et al., 2018).

Tendo isso em vista, a análise por imunoinformática de tumores de cavidade oral com alto e baixo potencial de metástase pode auxiliar a elucidar os mecanismos responsáveis pela metastização. A metástase desse tumor reduz consideravelmente a sobrevida dos pacientes e a identificação de marcadores de metastização ajudará na compreensão da biologia do tumor e, eventualmente, auxiliará na intervenção cirúrgica e tratamento.

2 Objetivo

2.1 Objetivo geral

Identificar indícios de evasão tumoral em tumores metastáticos e não metastáticos de cavidade oral HPV negativos com enfoque imunológico.

2.2 Objetivos específicos

- Relacionar genes diferencialmente expressos e vias celulares enriquecidas com a metastização tumoral;
- Associar a presença de populações imunológicas características de cada condição com a metastização;
- Caracterizar e comparar o perfil de mutações e neoantígenos entre as duas condições com o perfil metastático das mesmas;
- Identificar indícios de inibição imunológica e/ou escape tumoral bem como indícios de imunoeedição.

3 Material e Métodos

Neste estudo foram analisadas duas coortes. Na coorte do AC Camargo (ACC), tivemos acesso aos dados brutos que foram então pré-processados, diferentemente da segunda coorte, cujos dados do TCGA já se encontram pré-processados e seguiram as práticas do Broad Institute (BROAD INSTITUTE, 2018). A Tabela 3.1 sumariza as metodologias aplicadas a cada uma das coortes e a Figura 3.1 o fluxo de análise (*pipeline*) utilizado.

Tabela 3.1 Sumário das análises realizadas para cada coorte.

	ACC	TCGA
Pré-processamento dos dados	x	
Expressão diferencial e enriquecimento de vias	x	x
Variantes e predição de NeoAg		x
Identificação de TCR e BCR		x
Estimativa de populações do microambiente	x	x

Os dados brutos da coorte ACC foram pré-processados por Netto e colaboradores através do Lifescope versão 2.5.1 da Life Technologies e compartilhados em formato fastq. A partir do arquivo fastq foi realizada a avaliação da qualidade das leituras pelo FastQC e seleção das mesmas pelo Trimmomatic (Fig 3.1). O alinhamento contra o genoma humano foi então realizado com o alinhador star e a expressão dos genes com o pacote RSEM. A partir desses dados de expressão foram avaliados os genes diferencialmente expressos (DEG), as vias enriquecidas e o microambiente tumoral com o xCell. Para os dados disponíveis no TCGA foi também identificado o alelo de HLA de cada amostra pelo Optitype para subsequente predição de NeoAg com o netMHCpan. O repertório de receptores de linfócitos T (TCR) e B (BCR) foram identificados com o programa MiXCR.

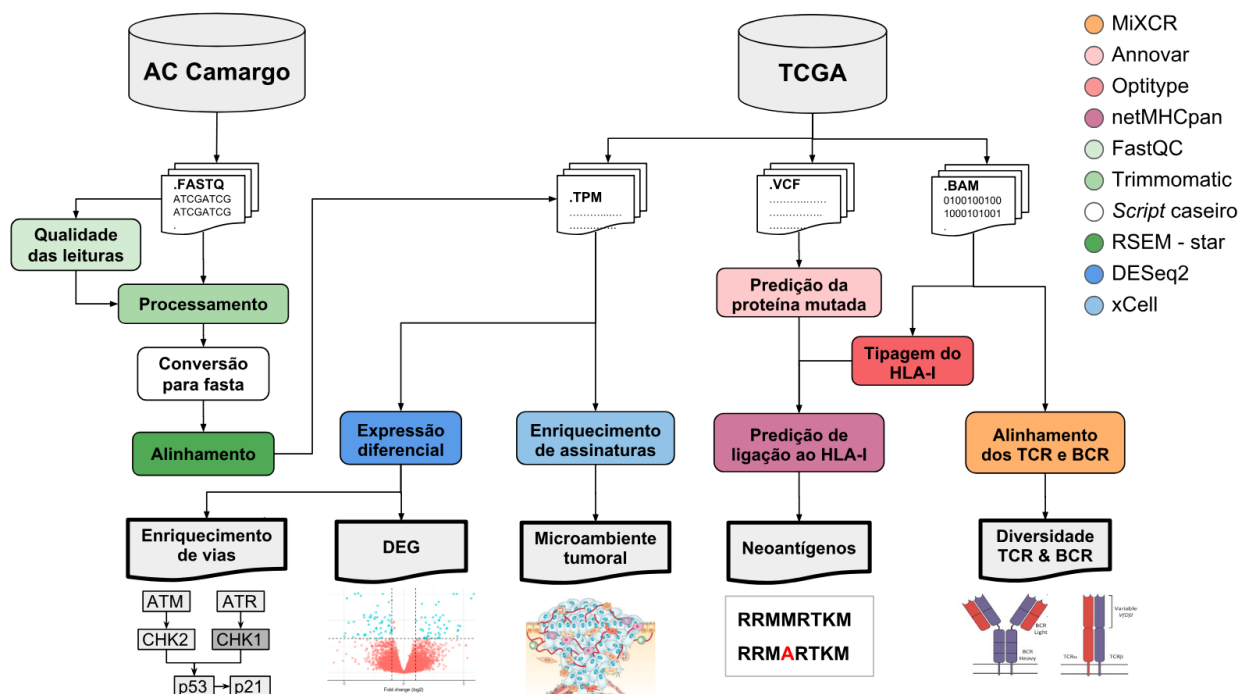


Figura 3.1 Esquematização do *pipeline* de análises desenvolvido pelos grupos e utilizado neste trabalho. A partir do tipo de dado, processado (TPM, vcf, bam) ou bruto (fastq), são aplicadas diferentes abordagens de processamento e análise. Os dados brutos da coorte do AC Camargo passaram por passos de verificação da qualidade e alinhamento ao genoma de referência antes de serem realizadas as análises específicas. O mesmo não ocorreu com os dados da coorte do TCGA, pois os mesmos já se encontram pré-processados. TPM - Transcrito por milhão. VCF - Formato de chamada de variante. BAM - Mapa de alinhamento binário. DG - Genes diferencialmente expressos. Figuras adaptadas de: (JOYCE; POLLARD, 2009; MURPHY, 2012g).

3.1 Obtenção das amostras

3.1.1 Coorte selecionada pelo AC Camargo

A coorte ACC, cedida por colaboradores, foi obtida a partir da seleção de amostras tumorais de cavidade oral (língua e assoalho bucal) HPV negativas classificadas patologicamente em t1 ou t2 com metástase linfonodal e classificadas em t3 ou t4 sem metástase linfonodal (NETTO, 2015). Os dados do sequenciamento foram gerados na plataforma SOLiD no modo paired-end (75 bp x 35 bp) a partir de amostras de cDNA. As amostras foram sequenciadas em duas corridas diferentes. Após a inclusão das amostras da coorte ACC, foram selecionadas 10 amostras sem metástase linfonodal e 10 com metástase linfonodal (Tabela 3.2). Cada grupo possui o mesmo número de amostras provenientes do sexo masculino (n=8) e feminino (n=2). Não foram incluídas na análise amostras em estadiamento t2 e o sítio anatômico predominante no grupo não metastático foi a língua (código C02 da Classificação Internacional de Doenças - ICD (WHO, 2016) enquanto no grupo

metastático 5 amostras estavam localizadas na língua e 5 no assoalho bucal (C04 - ICD). A idade média foi de 58,4 anos no grupo metastático e de 57,2 no grupo não metastático.

Tabela 3.2 Características das amostras da coorte do ACC

	Metástase linfonodal		Sem metástase linfonodal	
	Corrida 1	Corrida 2	Corrida 1	Corrida 2
Gênero				
Homem	5 (100%)	3 (60%)	5 (100%)	3 (60%)
Mulher	0	2 (40%)	0	2 (40%)
Sítio anatômico				
Assoalho bucal (C04)	5 (100%)	0	0	0
Língua (C02)	0	5 (100%)	5 (100%)	5 (100%)
Estadiamento patológico do tumor				
t1	5 (100%)	5 (100%)	0	0
t3	0	0	0	4 (80%)
t4	0	0	5 (100%)	1 (20%)
Estadiamento do linfonodo adjacente				
N1	2	2	0	0
N2b	2	3	0	0
N3	1	0	0	0

3.1.2 Critérios de inclusão da coorte do TCGA

As amostras foram selecionadas do repositório público *Tumor Cancer Genome Atlas* (TCGA), projeto TCGA-HNSC (*Head and neck squamous cancer*) que possui um total de 524 casos. Os casos foram filtrados para amostras HPV negativas, do assoalho da boca ou língua e sem tratamento neoadjuvante, que foram subsequentemente selecionados de modo a se obter tumores classificados patologicamente em t2 ou inferior com presença de metástase linfonodal ou aqueles classificados em t3 ou superior com ausência de metástase linfonodal.

3.2 Pré-processamento dos dados

A qualidade das leituras foi avaliada através do programa Fastqc (versão 0.11) (BABRAHAM BIOINFORMATICS, 2018). As leituras senso e antissenso foram separadas com um programa *in house*. O apuramento de bases de baixa qualidade e seleção das leituras de boa qualidade foi realizado com o programa Trimmomatic (versão 0.32) (BOLGER; LOHSE; USADEL, 2014) com parâmetros ILLUMINACLIP:2:30:10 SLIDINGWINDOW:4:10 TRAILING:15 MINLEN:30 e sequência de adaptadores (Adaptador 1: CGCCTTGGCCGTACAGCAG; Adaptador 2: CTGCCCCGGGTTCTCATTCT). O parâmetro ILLUMINACLIP exclui sequências de adaptadores aceitando até 2 pareamentos errôneos e com determinado score de alinhamento (30:10); o SLIDINGWINDOW realiza uma varredura nas leituras aceitando somente bases com média de qualidade na escala phred igual ou superior a 10 em uma janela de 4 bases; o parâmetro TRAILING exclui bases do final da leitura que estejam abaixo do valor passado (15); e o MINLEN filtra para leituras com tamanho igual ou superior a 30. A escala Phred é uma escala logarítmica da probabilidade de erro da chamada de base, e.g. um valor de 30 nessa escala equivale a 99,9% de certeza na chamada da base enquanto um valor de 40 equivale a 99,99% (EWING et al., 1998). As leituras que passaram por esse filtro foram agrupadas em um só arquivo e convertidas de arquivo fastq para arquivo fasta por programa *in house*. O programa Prinseq (versão 0.20) (SCHMIEDER; EDWARDS, 2011) foi executado com os parâmetros -trim_tail_left 7 -trim_tail_right 7 -min_len 30. O parâmetro trim_tail exclui bases poliA ou poliT nas extremidades da leitura contendo um número mínimo de bases determinado, no caso, 7; o min_len é

análogo ao MINLEN do Trimmomatic. O cálculo de leituras por transcrito foi realizado através do pacote RSEM versão 1.3.1 no modo single-end (LI; DEWEY, 2011) utilizando o alinhador star, parâmetro --no-qualities (que alinha arquivos fasta), genoma de referência GRCh38 versão 83 desconsiderando os pseudogenes e utilizando parâmetros padrão do programa (DOBIN, 2014).

3.3 Expressão diferencial e enriquecimento de vias

O pacote DESeq2 (versão 1.22.1) (LOVE; HUBER; ANDERS, 2014) do repositório do Bioconductor para R foi utilizado para determinação dos genes diferencialmente expressos entre a condição metastático e não metastático. Nas amostras da coorte do AC Camargo foi incluído um parâmetro que considera o efeito de lote (*batch effect*) dentro do desenho da matriz do DESeq2. Foram considerados, para ambas as coortes, genes diferencialmente expressos (DEG) aqueles com valor de *log2fold change* superior ou igual a 1 e valor de p ajustado menor ou igual a 0,05. As contagens normalizadas foram calculadas por rlog do pacote DESeq2. A análise de enriquecimento de vias foi realizada a partir dos 3000 genes com maior e menor valor de *log2fold change* de cada coorte, utilizando o método de *Gene Set Enrichment Analysis* (GSEA), por meio da ferramenta online WebGestalt. Três bancos de dados foram utilizados consultados em nossas análises: KEGG (*Kyoto Encyclopedia of Genes and Genomes*), PANTHER (*Protein ANalysis THrough Evolutionary Relationships*) e Reactome. As vias encontradas foram filtradas usando critérios de $FDR \leq 0,05$.

3.4 Carga de mutação e predição de ligação dos peptídeos ao HLA

De modo a aumentar a confiabilidade do resultado, foram analisadas as variantes identificadas por três programas diferentes (VarScan2, MuSE e SomaticSniper). As variantes de base única compartilhadas entre o VarScan2 e pelo menos um dos outros dois programas foram mantidas. As inserções e deleções são somente identificadas pelo programa VarScan2 e foram mantidas. Em seguida, as variantes somáticas foram selecionadas para manter somente aquelas que ocasionam mudança na sequência da proteína (i.e. polimorfismo de base única não sinônimo, e pequenas inserções e deleções). Foram consideradas somente mutações em genes mutados com valor de expressão na escala Transcritos Por

Milhão (TPM) superior ao quartil inferior desse gene nas amostras para a predição de NeoAg. Genes de expressão aberrante com TPM > 10 e sem expressão no tecido normal também foram selecionados para a predição de NeoAg derivados das suas proteínas correspondentes. A sequência de aminoácidos da proteína mutada e da correspondente selvagem foram geradas pelo ANNOVAR (versão 2018). Os alelos de HLA de classe I foram determinados com precisão de quatro dígitos para cada paciente utilizando o Optitype (versão 1.3.1) (SZOLEK et al., 2014). Em seguida, foi utilizado o netMHCpan versão 4.0 (JURTZ et al., 2017) para determinar a afinidade de ligação dos peptídeos à fenda do HLA-I. Os peptídeos ligados fortemente, com $IC_{50} \leq 50nM$, foram selecionados. Os antígenos exclusivos derivados das sequências mutadas foram ainda filtrados em um banco de NeoAg gerados pelo netMHCpan a partir da sequência referência e os antígenos que superaram o filtro foram chamados de NeoAg. Peptídeos com ligação predita a mais de um alelo de HLA do mesmo paciente foram contados uma única vez. As mutações em genes de reparo e da via de processamento e apresentação de antígenos foram pesquisados conforme os Anexo 1 e 2.

3.5 Identificação de TCR e BCR

A identificação de receptores de linfócito T e B foi feita utilizando o programa MiXCR v. 1.3.1 (BOLOTIN et al., 2017) a partir dos dados brutos de RNASeq das amostras do TCGA. Os parâmetros utilizados foram os recomendados pelos desenvolvedores e foram contabilizadas somente sequências sem códon de parada prematuro e sem saída da fase de leitura. A análise dos receptores foi realizada com o pacote tcR (NAZAROV et al., 2015). A identificação de clones de TCR e BCR por meio do programa MiXCR é baseada no alinhamento de regiões constantes FR3 e 4 (FR3, *Framework region* 3) adjacentes ao CDR3 (Região determinante de complementariedade 3). Esse alinhamento considera os rearranjos gênicos nos *loci* de TCR e BCR no caso de sequenciamento por RNA-Seq e também permite emparelhamentos errados de base, conforme configuração do programa.

3.6 Estimativa das populações do microambiente tumoral

As assinaturas de células do microambiente tumoral (estromais ou imunológicas) foram estimadas com o programa xCell (ARAN; HU; BUTTE, 2017) a

partir dos dados de expressão normalizados em TPM para a coorte ACC. Foram somente consideradas as estimativas com p-valor menor ou igual a 0,05. O dado para a coorte TCGA já estava disponível no repositório do xCell e foi utilizado. A partir dos valores das estimativas, foi realizada a correlação de Pearson substituindo aquelas com p valor $\geq 0,05$ ou número de observações ≤ 5 por 0; foram mostradas somente as correlações com o desfecho que possuem valor igual ou superior a 0,4. O programa CIBERSORT foi utilizado para confirmar algumas das assinaturas encontradas pelo xCell. Foram consideradas as estimativas com valor de $p \leq 0,05$ (NEWMAN et al., 2015). A assinatura de Treg e MDSC foram estimadas com uma assinatura de células do SI - *Immunophenoscore* (CHAROENTONG et al., 2017).

3.7 Análises estatísticas e geração de dados em R

Os gráficos e pacotes foram executados na versão 3.5.1 do R e versão 1.1.456 do RStudio. Os principais pacotes utilizados foram o DESeq2 versão 1.20.0, pheatmap versão 1.0.10, ggplot2 versão 3.1 e tcR versão 2.2.2, tximport 1.10.1.

4 Resultados

4.1 Características das amostras selecionadas

Na coorte do TCGA, foram selecionadas 11 amostras no grupo de metástase e 11 no grupo sem metástase linfonodal (Tabela 4.1). Os grupos são formados por amostras de pacientes do sexo masculino estando igualmente distribuídas entre o assoalho bucal (n=3) e a língua (n=8). Ao contrário da coorte ACC, 10 de 11 amostras do grupo metastático são formadas por tumores no estadiamento t2. Não houve diferença significativa ($p < 0,05$ no teste exato de Fisher) entre os grupos para gênero, sítio anatômico, consumo de álcool e tabagismo.

Tabela 4.1. Características das amostras da coorte do TCGA.

	Metástase linfonodal	Sem metástase linfonodal
Gênero		
Homem	11	11
Mulher	0	0
Sítio anatômico		
Assoalho bucal (C04)	3 (27,3%)	3 (27,3%)
Língua (C02)	8 (72,7%)	8 (72,7%)
Estadiamento patológico do tumor		
t1	1 (9,1%)	0
t2	10 (90,9%)	0
t3	0	4 (36,4%)
t4a	0	7 (63,6%)
Consumo de álcool		
Sim	8 (72,7%)	5 (45,5%)
Não	3 (27,3%)	6 (55,5%)
Tabagismo		
Fumante	8 (72,7%)	5 (45,4%)
Ex-fumante (até 15 anos)	2 (18,2%)	4 (36,4%)
Ex-fumante (mais de 15 anos)	1 (9,1%)	1 (9,1%)
Ex-fumante (duração não informada)	0	1 (9,1%)

4.2 Pré-processamento dos dados

4.2.1 Processamento e filtragem dos dados brutos

As 20 amostras provenientes da colaboração com o AC Camargo foram sequenciadas em duas corridas. A Figura 4.1 apresenta um boxplot com a distribuição dos valores de qualidade na escala Phred para cada base na leitura (read). Durante a corrida 2, houve uma parada precoce por volta da base 50 e retomada da corrida até a base 68, ao invés da base 75. Isso impactou na qualidade das leituras entre a parada e retomada que ficou muito abaixo da recomendada (zona vermelha, Figura 4.1-A e Figura 4.2-A). A corrida 1 decorreu sem problemas, a qualidade das bases ficou em sua maioria nas regiões verde e amarela que possuem um valor de qualidade na escala Phred mais elevado (Fig 4.1-D). Após o apuramento e filtragem das sequências, a qualidade das leituras em geral ficou acima de 28 na escala Phred para a corrida 1 e de Phred 20 para a corrida 2 (Fig 4.1-B). A distribuição dos tamanhos das leituras é melhor visualizados na Figura 4.2-A que apresenta dois picos indicando os tamanhos de sequências (50 e 68) decorrentes da parada e retomada da corrida 2. Também após a filtragem foi observada a prevalência de sequências em torno de 50bp para a corrida 2 enquanto (Fig 4.2-C) na corrida 1 predominou o mesmo tamanho de leitura pré-filtragem de qualidade - 75bp (Fig 4.2-B e 4.2-D). Não foram detectadas pelo FastQC sequências super-representadas ou de adaptadores do sequenciamento.

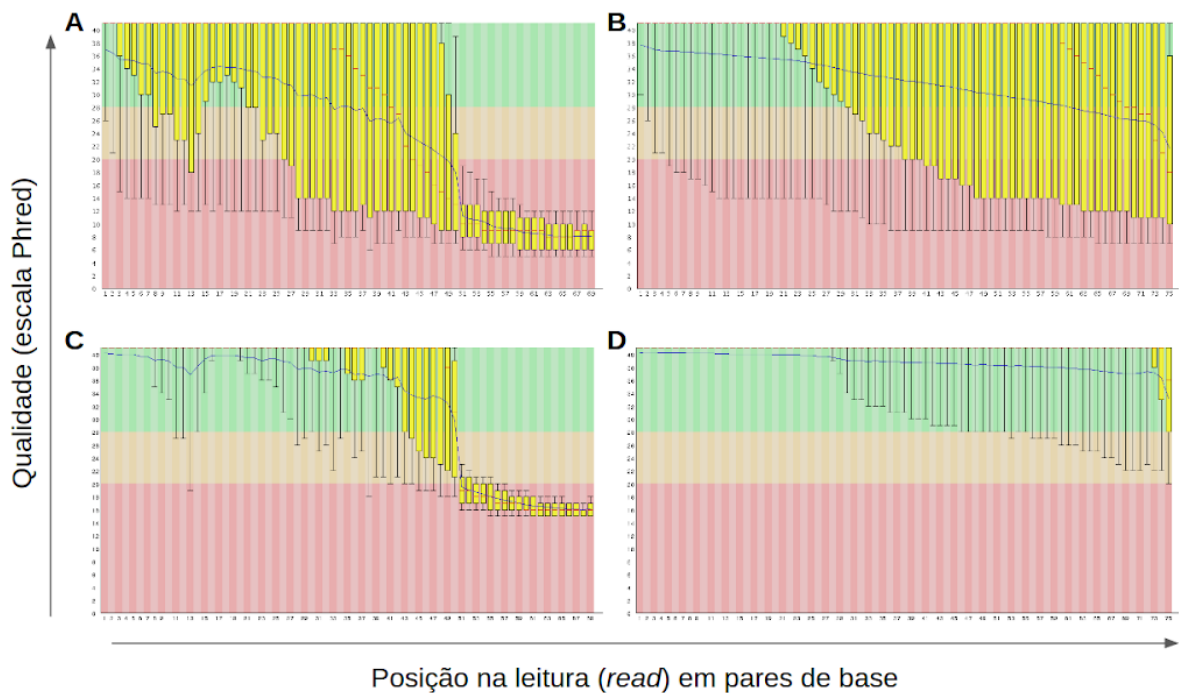


Figura 4.1 Qualidade das leituras da coorte ACC geradas pelo FastQC. A qualidade está na escala Phred sob a forma de boxplot representada para cada base do sequenciamento. A barra vermelha dentro do boxplot representa a mediana e a linha azul contínua acompanha a tendência da dispersão. Amostra da qualidade de um arquivo bruto da corrida 2 (**A**) e corrida 1 (**B**). Mesmas amostras após exclusão de leituras de baixa qualidade (**C e D**).

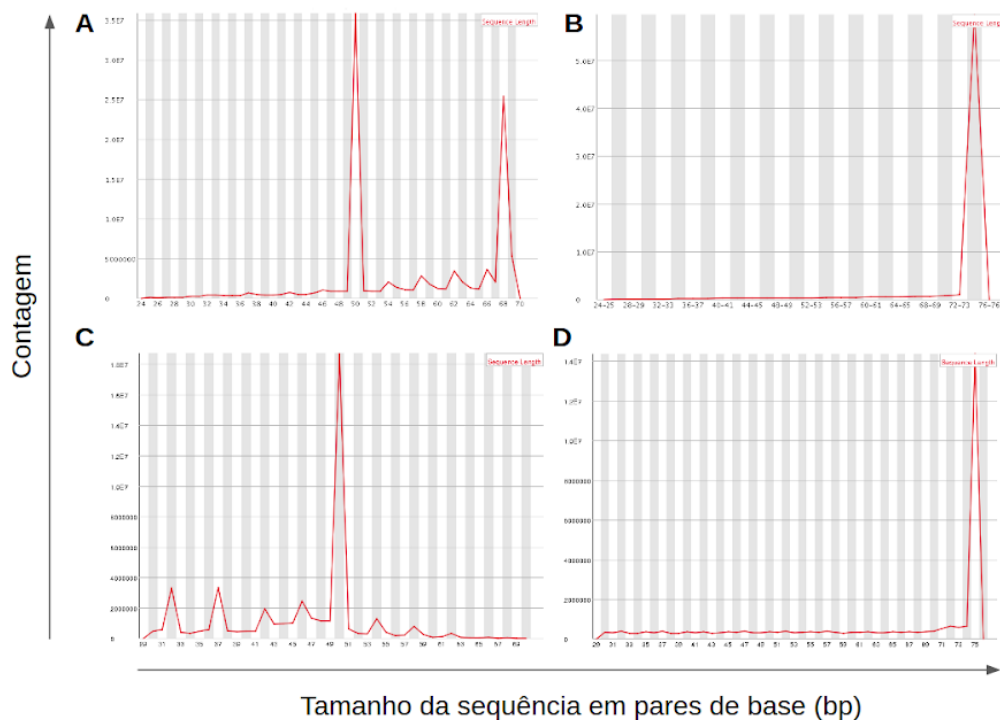
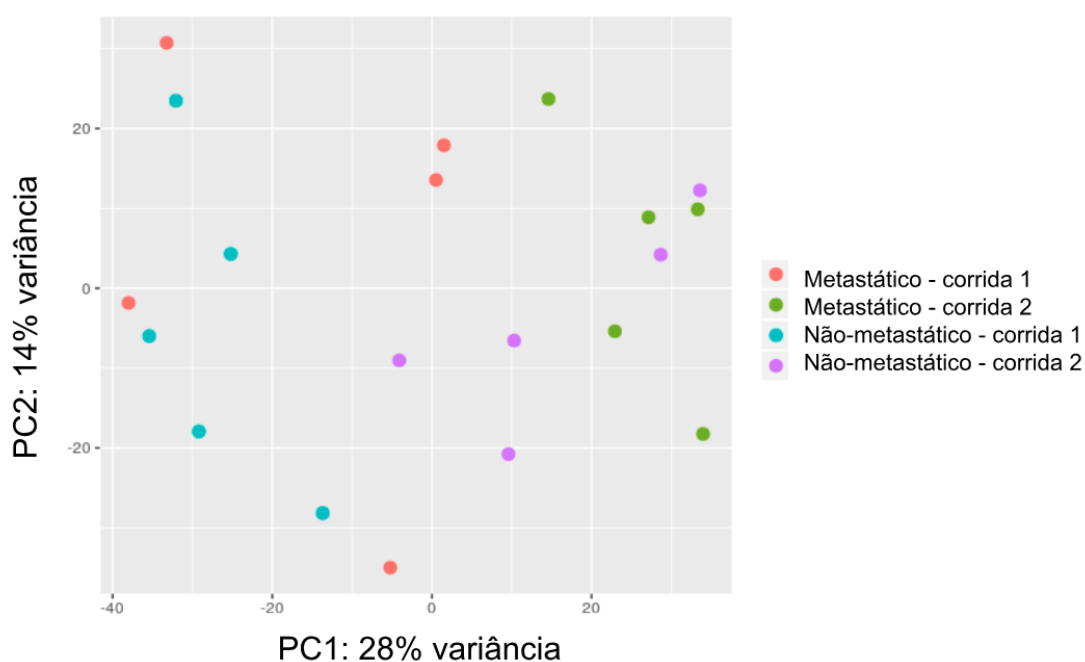


Figura 4.2 Distribuição do tamanho das sequências entre as leituras. Mesmas amostras brutas da corrida 2 (**A**) e 1 (**B**) mostradas na Figura 4.1. Mesmas amostras após exclusão de leituras com baixa qualidade para corrida 2 (**C**) e 1 (**D**).

4.2.2 Efeito de lote da coorte ACC

Uma análise de componente principal (PCA, *Principal Component Analysis*) realizada na coorte ACC demonstrou o viés introduzido pela diferença entre as corridas 1 e 2 (Fig 4.3). Nela é possível observar que os pontos azuis e rosa (corrida 1) se agrupam, assim como os pontos verdes e roxos (corrida 2). Por outro lado, as amostras do grupo metastático (corrida 1 e 2) não se agrupam tanto como as amostras de uma mesma corrida. Ou seja, a maior variância entre as amostras (PC1) evidencia a separação entre corridas e não entre desfecho (metastático e não metastático).



dispersão gerado pelo programa, no qual os pontos pretos representam os valores brutos e servem de base para o traçado da curva de tendência - em vermelho. Em seguida, os valores de dispersão (pontos pretos) são normalizados por essa curva, resultando nos pontos azuis. Alguns genes falham nesse processo de normalização e são representados como pontos pretos com um contorno azul. Ambas as coortes tiveram uma curva de tendência que acompanhou a dispersão da maioria dos genes. A coorte ACC apresentou uma quantidade maior de genes com valores *outlier* de dispersão, representados pela nuvem de pontos pretos com contorno azul (Fig 4.4, superior).

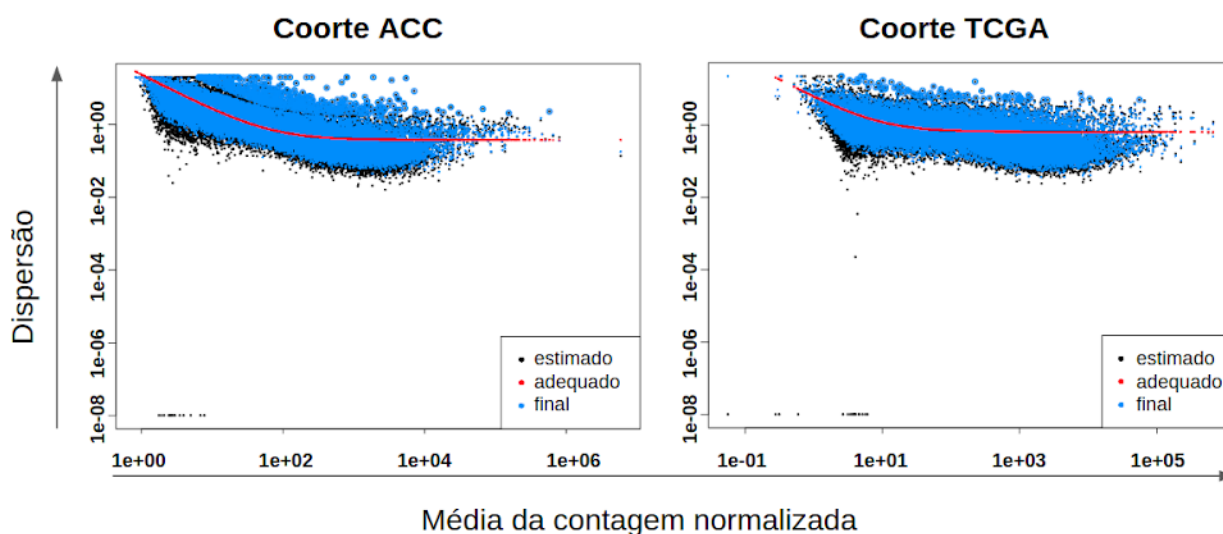


Figura 4.4 Estimativa de redução da dispersão das contagens normalizadas pelo DESeq2 nas duas coortes. Os pontos pretos representam valores brutos estimados de dispersão gênica; a curva vermelha é a média de dispersão entre genes com uma mesma contagem; e os pontos azuis os valores após a normalização dos valores brutos pela curva de tendência (vermelha). Pontos pretos com contorno azul são valores *outlier* de dispersão.

4.3 Genes diferencialmente expressos e vias enriquecidas

Foi realizada uma análise dos genes diferencialmente expressos (DEG) entre as condições estudadas a fim de se identificar genes associados com o perfil metastático dos tumores. Utilizando um valor de corte de p ajustado $\leq 0,05$ e de $\log_2 \text{fold change} \geq 1$, foram identificados 186 DEG na coorte do TCGA, 127 na coorte do ACC. A fim de melhor visualizar quais são os DEG em cada coorte, foi plotado um Volcano plot que distribui os genes por variação da expressão entre os grupos ($\log_2 \text{fold change}$) e valor de p ajustado. Em ambas as coortes, os genes em azul estão com maior expressão diferencial na condição não metastática e os em vermelho na condição metastática; os genes em cinza estão abaixo do valor de corte

citado. A coorte ACC possui genes da família *MAGE* (*MAGEA12*, *MAGEA2* e *MAGEA2B*) entre os mais diferencialmente expressos do grupo não metastático, assim como a Interleucina 1 alfa (*IL1A*) conforme mostra a Figura 4.5-A - lado esquerdo. No grupo metastático, um gene de destaque é o fator de crescimento de fibroblasto 19 (*FGF19*) - Figura 4.5-A - lado direito. O grupo não metastático da coorte TCGA possui o *CD300E*, um receptor de células mieloides, como um dos genes mais diferencialmente expressos (Fig 4.5-B - lado esquerdo). Já o grupo metastático possui a *ACTA1* e *ACTN2* (Alfa Actina 1 e 2, respectivamente) como alguns dos genes mais diferencialmente expressos (Fig 4.5-B - lado direito).

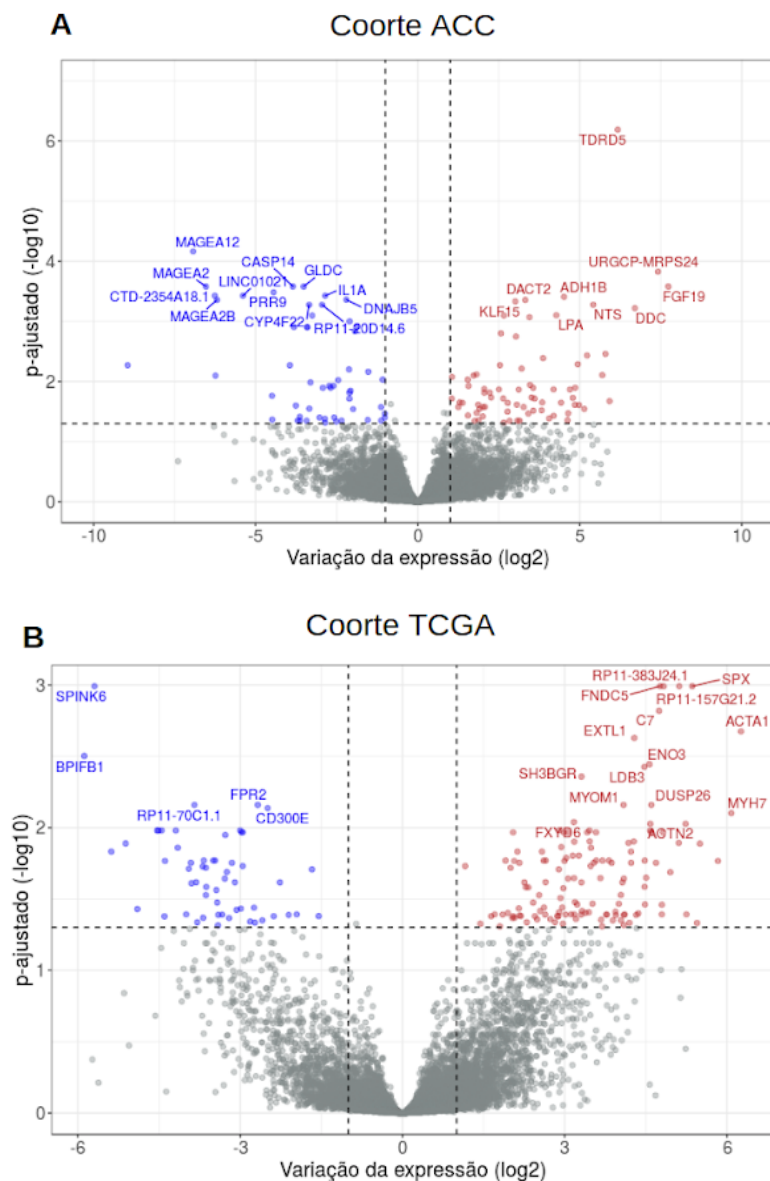


Figura 4.5 Volcano plot dos genes diferencialmente expressos entre a condição com metástase e sem metástase para a coorte do ACC **(A)** e do TCGA **(B)**. Em azul os genes mais diferencialmente expressos na condição não metastática e em vermelho os na condição metastática. Genes em cinza não passaram no filtro (p ajustado $\leq 0,05$ e $\log_2\text{fold change} \geq 1$). Os 22 genes com menor p-valor ajustado estão destacados.

Quando comparamos os DEGs entre as duas coortes, encontramos apenas 3 genes em comum (Fig 4.6). Dos 186 DEG exclusivos à coorte do TCGA, 133 estão mais expressos na condição metastática e 56 na não metastática. Já dos 127 DEG da coorte ACC, 75 estão mais expressos na condição metastática e 55 na não metastática. O gene *PIWIL2* (ENSG00000197181) está mais expresso na condição não metastática em relação à metastática (ACC, p-ajustado=0.0419, contagem média normalizada: 112.497; TCGA, p-ajustado=0.0191, contagem média normalizada: 348.4564) e o gene *ADH1B* (ENSG00000196616) está mais expresso na condição metastática em relação à não metastática (ACC, p-ajustado=0.000390, contagem média normalizada: 60.19857; TCGA, p-ajustado=0.04069092, contagem média normalizada: 77.15721). O terceiro gene, *C10orf105*, apesar de diferencialmente expresso, não foi concordante entre as coortes.

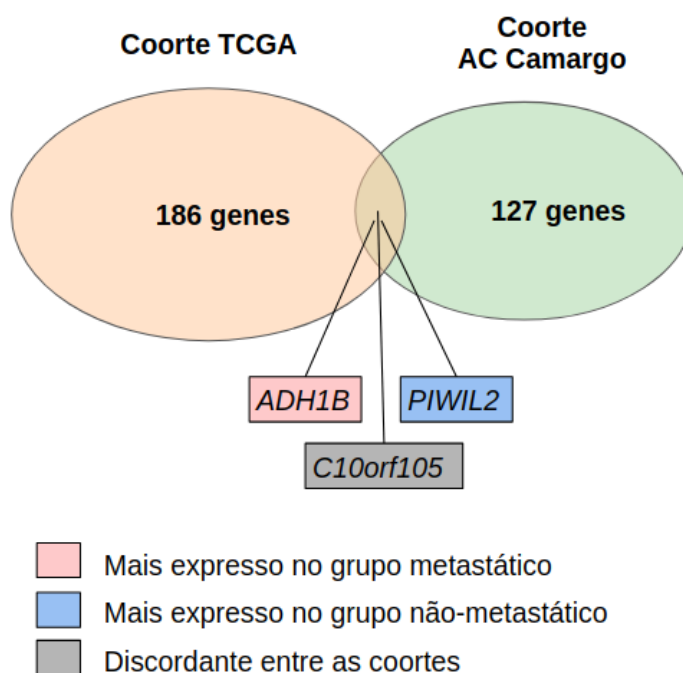


Figura 4.6 Genes diferencialmente expressos entre as duas coortes mostraram o gene *ADH1B* (álcool desidrogenase subunidade beta) e o gene *PIWIL2* (Piwi Like RNA-Mediated Gene Silencing 2) como os únicos concordantes entre as duas coortes. O *C10orf105* está mais expresso no grupo metastático na coorte ACC.

Quando comparada a expressão normalizada de *PIWIL2*, *C10orf105* e *ADH1B* entre os grupos, a *ADH1B* e *C10orf105* tem diferença significativa ($p < 0,01$, comparação de médias pelo método Wilcoxon) na coorte ACC (Fig 4.7-A), mas não o *PIWIL2* ($p = 0,064$, teste de comparação de médias de Wilcoxon). Utilizando o mesmo teste na coorte do TCGA, somente o *C10orf105* apresentou diferença

estatisticamente significativa (Fig 4.7-B; $p=0,17$ para *ADH1B* e $p=0,51$ para *PIWIL2*). Os valores de *log2fold change* para *ADH1B*, *PIWIL2* e *C10orf105* na coorte ACC foi de 4,51, -2,86 e 3,51, respectivamente, e de 3,78, -3,63 e -2,95 na coorte TCGA.

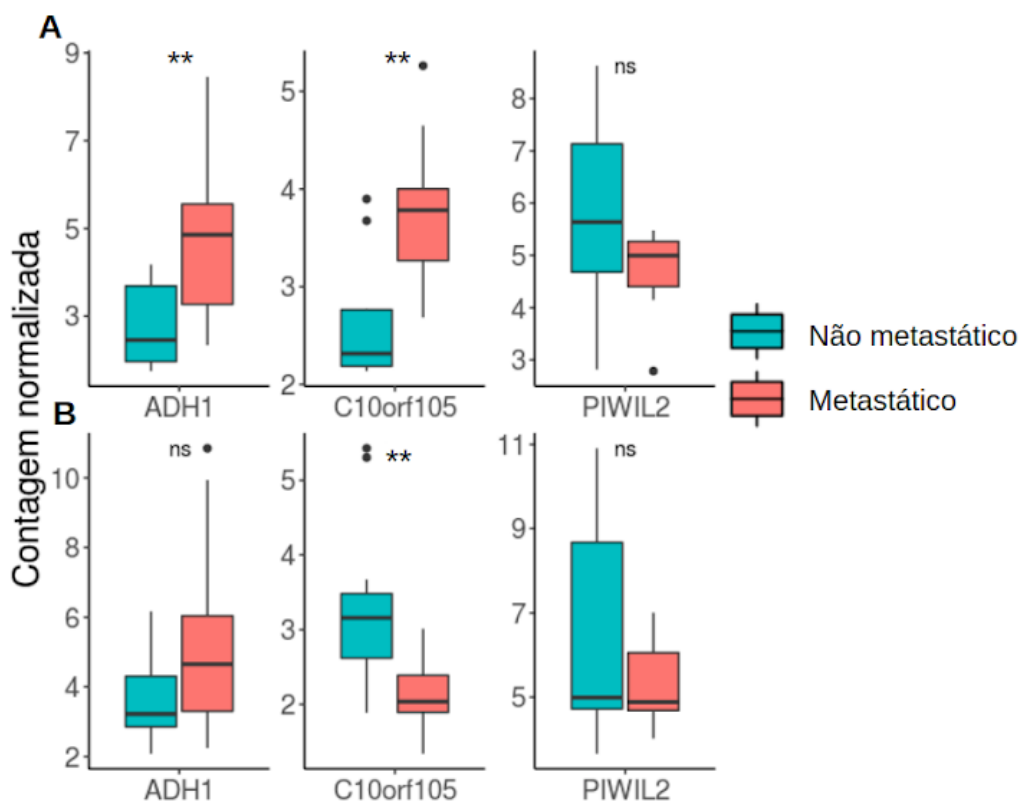


Figura 4.7 Contagem normalizada pelo DESeq2 (rlog) dos genes *ADH1B*, *C10orf105* e *PIWIL2* para a coorte ACC (A) e TCGA (B). Foi aplicado o teste de comparação de médias de Wilcoxon (Mann-Whitney) em todos os casos. ** $p < 0,01$.

O padrão de expressão dos DEG está representado sob a forma de um mapa de calor utilizando distância euclidiana para agrupamento de colunas (Fig 4.8). Na coorte ACC 7 amostras do grupo metastático formam um *cluster* bem definido à esquerda e 3 amostras se misturam entre as amostras do grupo não metastático (Fig 4.8-A). Já na coorte do TCGA, os grupos possuem uma separação mais heterogênea conforme mostra o dendrograma da Figura 4.8-B. Três *clusters* ficam evidentes na Figura 4.8-B, um formado por 8 amostras do grupo não metastático e dois formados por amostras de ambos os grupos. Em ambas as coortes não há uma separação total dos grupos.

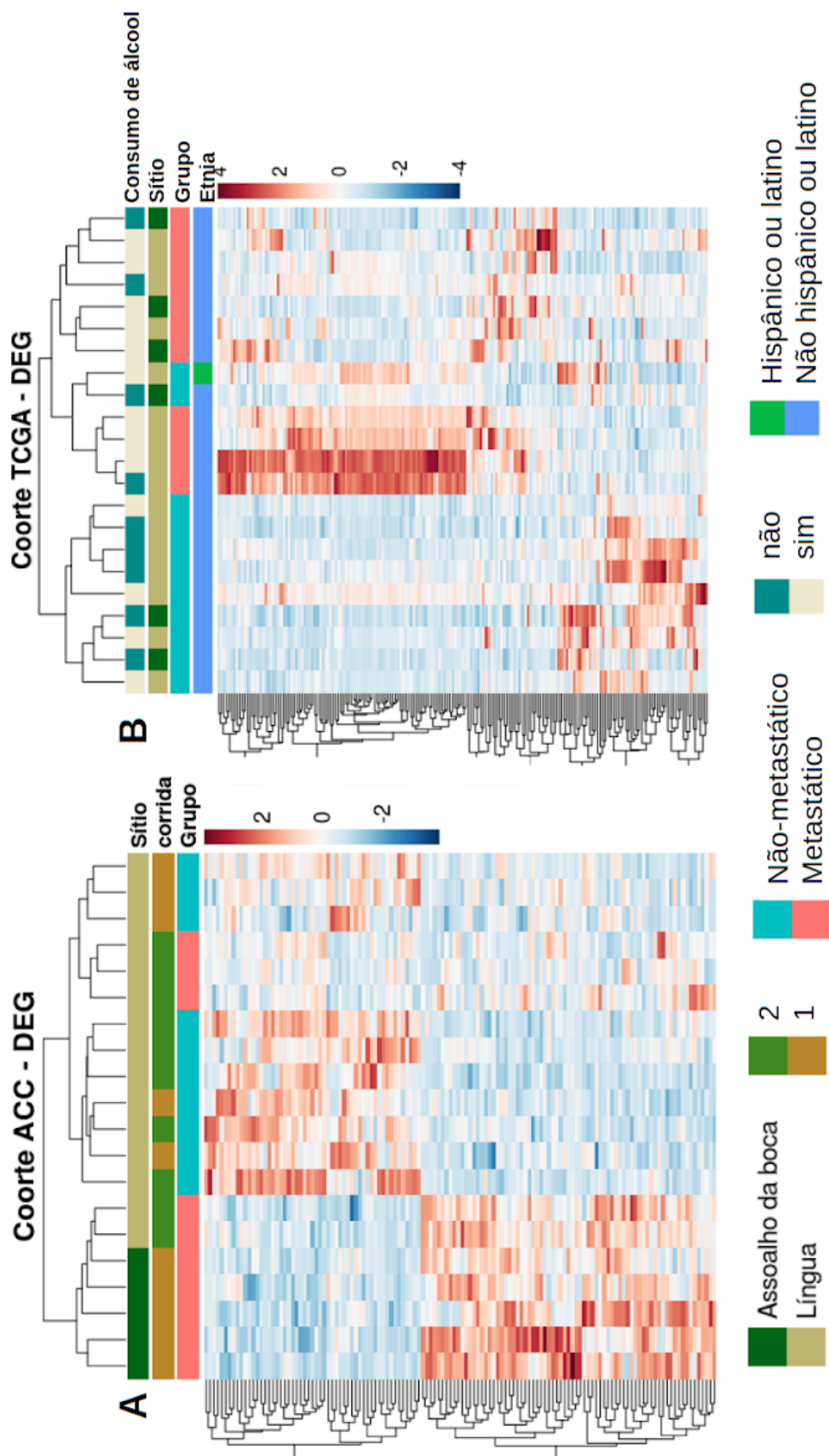


Figura 4.8 Mapa de calor dos genes diferencialmente expressos nas duas coortes (DEG). **A.** DEG na coorte ACC. **B.** DEG na coorte TCGA. Em ambos foi utilizada distância euclidiana para o agrupamento das colunas. São indicadas nas colunas alguns parâmetros clínicos e metodológico tais como o sítio anatômico, a corrida de sequenciamento, o grupo, consumo de álcool e etnia.

Genes não diferencialmente expressos podem estar relacionados ao fenótipo tumoral, ainda que de maneira não tão fortemente associada ou isolada. A identificação dos DEG com o pacote DESeq2 permitiu o cálculo do *log2fold change* para cada gene em relação à condição do tumor - metastático ou não. A partir desse dado, foi possível ranquear os genes por maior valor de módulo de *log2fold change* e realizar análises de enriquecimento de vias por *Gene Set Enrichment Analysis* (GSEA) nas quais pequenas variações da expressão desses genes poderiam estar associadas à maior ou menor representatividade de uma via celular. As vias foram selecionadas pela taxa de descoberta de falsos positivos ($FDR \leq 0,1$). Foram encontradas 3 vias enriquecidas no banco de dados PANTHER, 13 vias no KEGG e 11 vias no Reactome, sendo que todas as vias com *score* positivo de enriquecimento pertencem à coorte do TCGA (Fig 4.9). Dentre as vias mostradas, uma via enriquecida no grupo metastático está empobrecida no não metastático e uma via enriquecida no grupo não metastático está empobrecida no metastático. Duas vias estão reguladas positivamente no grupo não metastático em ambas as coortes: a formação do envelope cornificado e queratinização. Em 6 das vias enriquecidas no grupo metastático da coorte TCGA encontra-se alguma associação com a contração muscular ou músculo cardíaco. O mesmo grupo possui vias importantes reguladas negativamente, tais como a sinalização por interferon, processamento e apresentação de antígeno e citotoxicidade mediada por célula NK. A coorte ACC possui vias relacionadas ao ribossomo e à degranulação de neutrófilo reguladas negativamente no grupo metastático e positivamente no metastático.



Figura 4.9 Análise de enriquecimento de vias para as duas coortes considerando diferentes bancos de dados (KEGG, Reactome e PANTHER). Vias enriquecidas no grupo metastático (superior) e não metastático (inferior). As vias foram selecionadas considerando $FDR \leq 0,01$. Em verde a coorte do ACC e em preto a do TCGA. Um valor positivo de *Score* normalizado do enriquecimento significa que a via está enriquecida e um valor negativo significa que está empobrecida.

4.4 Populações do microambiente tumoral

A fim de estimar as populações celulares no microambiente tumoral, foi realizada uma análise de enriquecimento de assinaturas de populações de 64 tipos celulares nas duas coortes. A Figura 4.10 mostra a quantificação das populações que possuem ao menos metade das amostras com estimativas não nulas de enriquecimento para cada população. As colunas estão agrupadas por distância euclidiana e a escala foi feita por linha. Parece haver uma separação parcial entre os grupos metastáticos e não metastáticos na coorte TCGA (Fig 4.10-B), mas não na coorte ACC (Fig 4.10-A).

Na tentativa de relacionar a presença das subpopulações com desfechos clínico-patológicos foi realizada a correlação de Pearson entre as estimativas de enriquecimento de células do microambiente tumoral e o desfecho de metástase linfonodal. Somente coeficientes de correlação com módulo ≥ 0.4 são mostrados. As correlações com valor 0 correspondem àquelas com valor de p menor que 0,05 ou menos de 5 observações para a correlação; os quadrados azuis mostram correlações com o grupo não metastático e os vermelhos com o metastático (Fig 4.11). Os linfócitos T CD4 de memória tiveram correlação negativa com o desfecho de metástase em ambas as coortes, com correlação de 0,78 na coorte ACC e de 0,71 na do TCGA. Outras populações estão relacionadas ao desfecho, mas somente a população dos linfócitos T CD4 de memória foi comum a ambas as coortes. Os Pró-linfócitos B tiveram uma associação de 0,77 com o grupo metastático (coorte TCGA). Utilizando outra metodologia, o CIBERSORT, foi também possível observar uma fração maior de linfócitos B de memória no grupo metastático ($p < 0,05$ na comparação de médias pelo teste Wilcoxon; dado não mostrado). As outras assinaturas não foram estatisticamente diferentes no CIBERSORT.

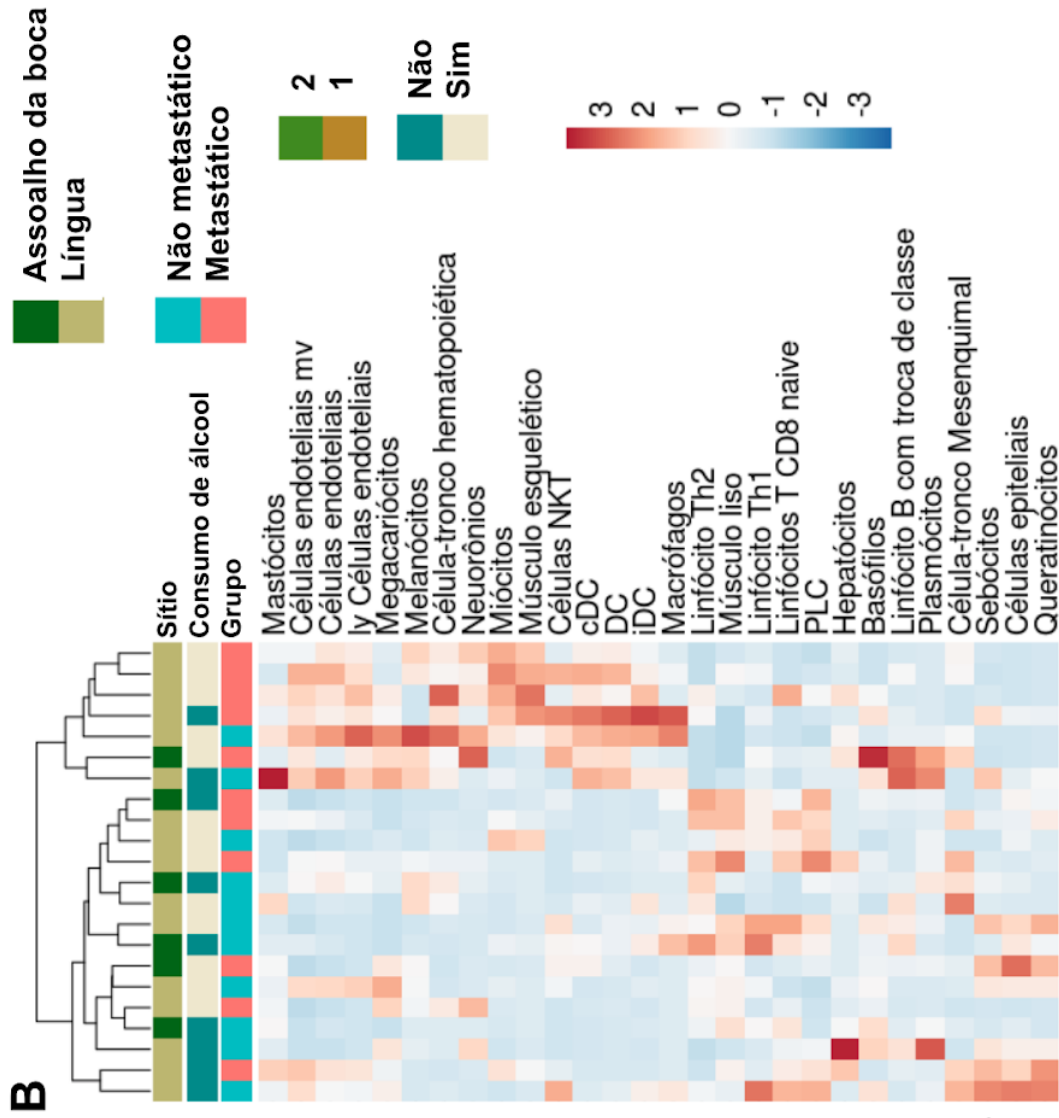
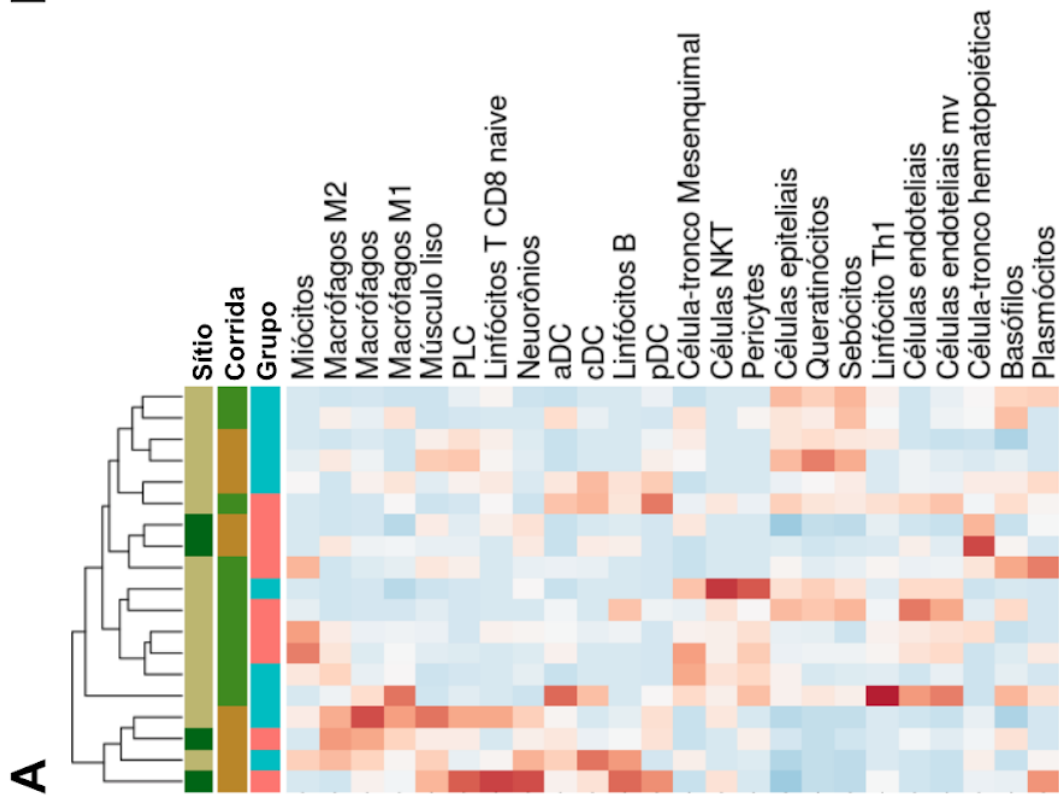


Figura 4.10 Mapa de calor da análise de enriquecimento de população para a coorte ACC **(A)** e TCGA **(B)**. A escala foi feita por linha e as colunas foram agrupadas pelo método de distância *manhattan*. **A.** Foi apresentada informação referente ao sítio anatômico, corrida do sequenciamento e grupos. **B.** Informações sobre sítio anatômico, consumo de álcool e grupo estão apresentadas na parte superior..

A grande maioria das assinaturas imunológicas com correlação acima de 0,4 estão associadas classe positivamente com o desfecho não metastático, e aquelas associadas com a metástase estão distribuídas entre populações imunológicas (linfócito Th1 e pró-linfócitos B) e estromais (músculo esquelético) ou células-tronco (Fig 4.11-A e 4.11-B). Foi realizada uma análise comparando os tumores HPV negativos de grau patológico t1 e t2 com aqueles t3 e t4 de modo a identificar enriquecimentos de populações associados ao tamanho tumoral, mas não foram observadas assinaturas com correlação de Pearson superior a 0,4.

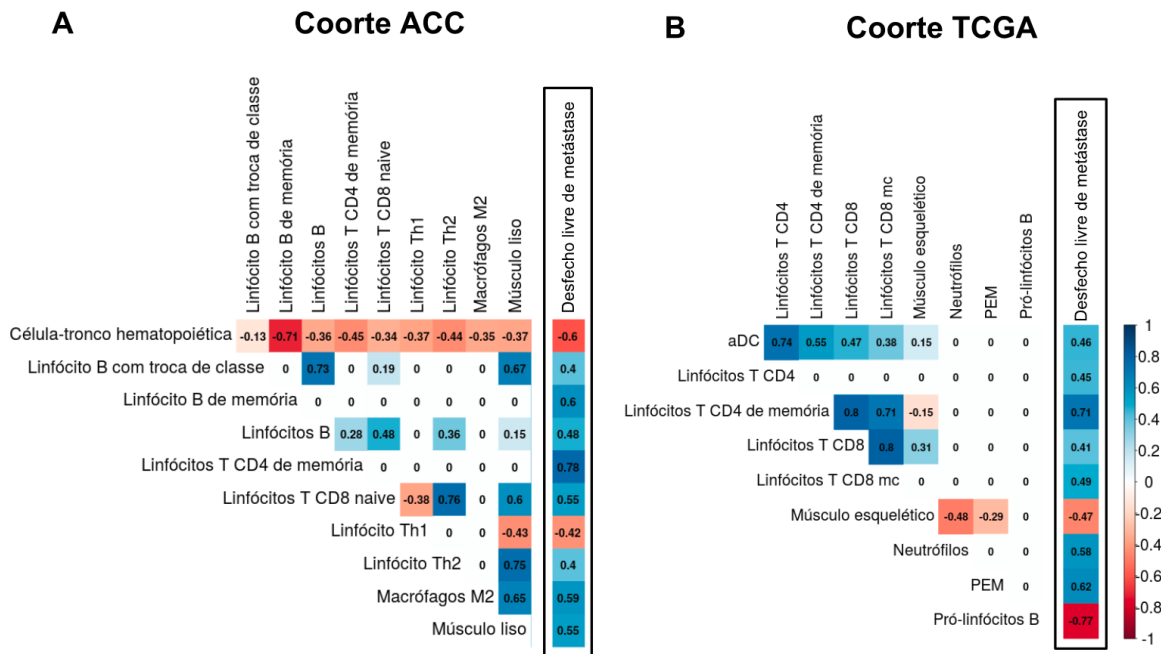


Figura 4.11 Correlações de Pearson entre as populações do microambiente e entre os grupos metastático e não metastático para a coorte ACC **(A)** e do TCGA **(B)**. A cor azul representa correlações positivas e, na última coluna, correlações favoráveis ao desfecho livre de metástase. Correlações com valor de $p > 0,05$ receberam o valor 0 de correlação. aDC: Célula dendrítica ativada. Linfócitos T CD8 mc: Linfócitos T CD8 de memória central. PEM - Progenitor eritroide-megacariócito.

4.5 Repertório de TCR e BCR

Foi possível identificar clones de TCR em 20 das 22 amostras analisadas sendo que as duas amostras sem clones de TCR detectável pertencem a grupos diferentes. A maioria dos clones são de cadeia alfa (47 no grupo não metastático e

47 no metastático) e beta (69 no grupo não metastático e 50 no metastático), mas também 1 a 2 clones de cadeia gama e delta (dado não mostrado) foram encontrados em 3 amostras do grupo não metastático (Fig 4.12-A). A diferença no número de clones de cadeia alfa e beta entre os grupos não foi significativa no teste de comparação de médias pelo método Wilcoxon (Mann-Whitney). O índice de diversidade foi calculado pelo pacote tcR usando o método *inverse Simpson* e não mostrou diferença significativa entre os grupos (Fig 4.12-B). Apesar disso, a representatividade clonal, uma medida indireta da expansão de um clone, é maior nos clones do grupo não metastático, no qual foram detectados 10 clones com pelo menos 3 leituras (*reads*) enquanto no grupo metastático foram detectados 3 clones (Fig 4.12-C).

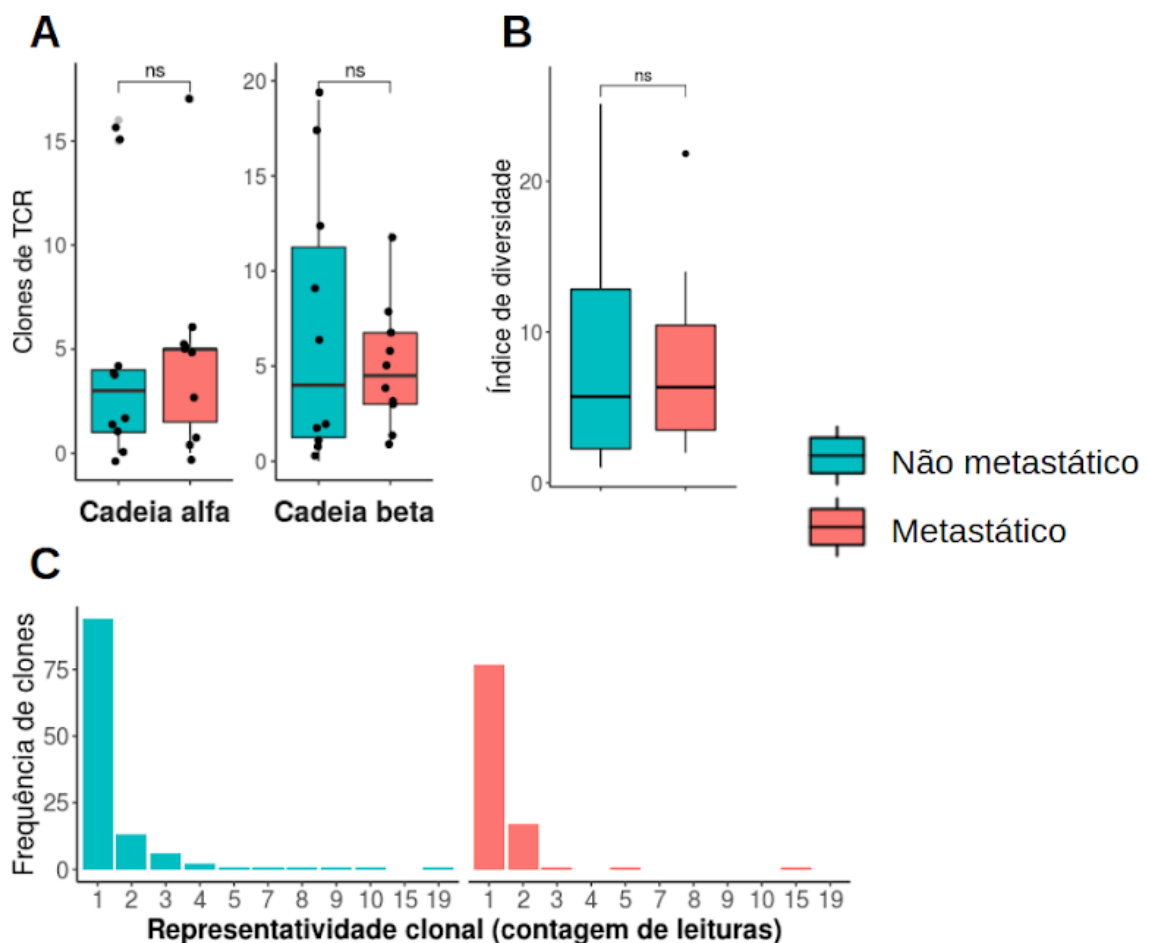


Figura 4.12 Repertório de TCR entre os grupos. **A.** A comparação do número de clones de TCR - cadeia alfa e beta - identificados mostrou que o grupo não metastático e o grupo metastático não possuem diferença significativa no teste de comparação de médias pelo método Wilcoxon (Mann-Whitney). **B.** Cálculo do índice de diversidade *inverse Simpson* não evidenciou diferença entre os grupos. **C.** Frequência da representatividade clonal. Cada clone possui um número de leituras (*reads*) que suporta sua identificação e serve como uma medida indireta da expansão desse clone.

Foi possível identificar clones de BCR em todas as amostras, contudo, algumas não contavam com a identificação da porção constante que caracteriza o subtipo de imunoglobulina (Ig). Considerando as porções de cadeia leve identificadas, kappa (total de 2.539) e lambda (total de 1.190), não houve diferença significativa entre os grupos pelo teste de comparação de médias do método Wilcoxon (Mann-Whitney; dado não mostrado; $p=0,69$ para cadeia kappa e $p=0,14$ para cadeia lambda). Dentre as cadeias pesadas identificadas, não foram encontrados os subtipos IgE e IgD, mas sim 12 clones de IgM (3 no grupo não metastático e 9 no metastático), 1.027 de IgG (803 e 224) e 225 de IgA (191 e 34), Figura 4.13-A. Não houve diferença estatística no índice de diversidade (Fig 4.13-B). Os clones com maior representatividade clonal estão presentes no grupo metastático (Fig 4.13-C). Uma amostra *outlier* do grupo não metastático que contava com 701 clones de IgG, 177 de IgA e 3 de IgM não foi representada nas Figuras 21-A e 21-C, mas foi considerada no cálculo das médias em 21-A.

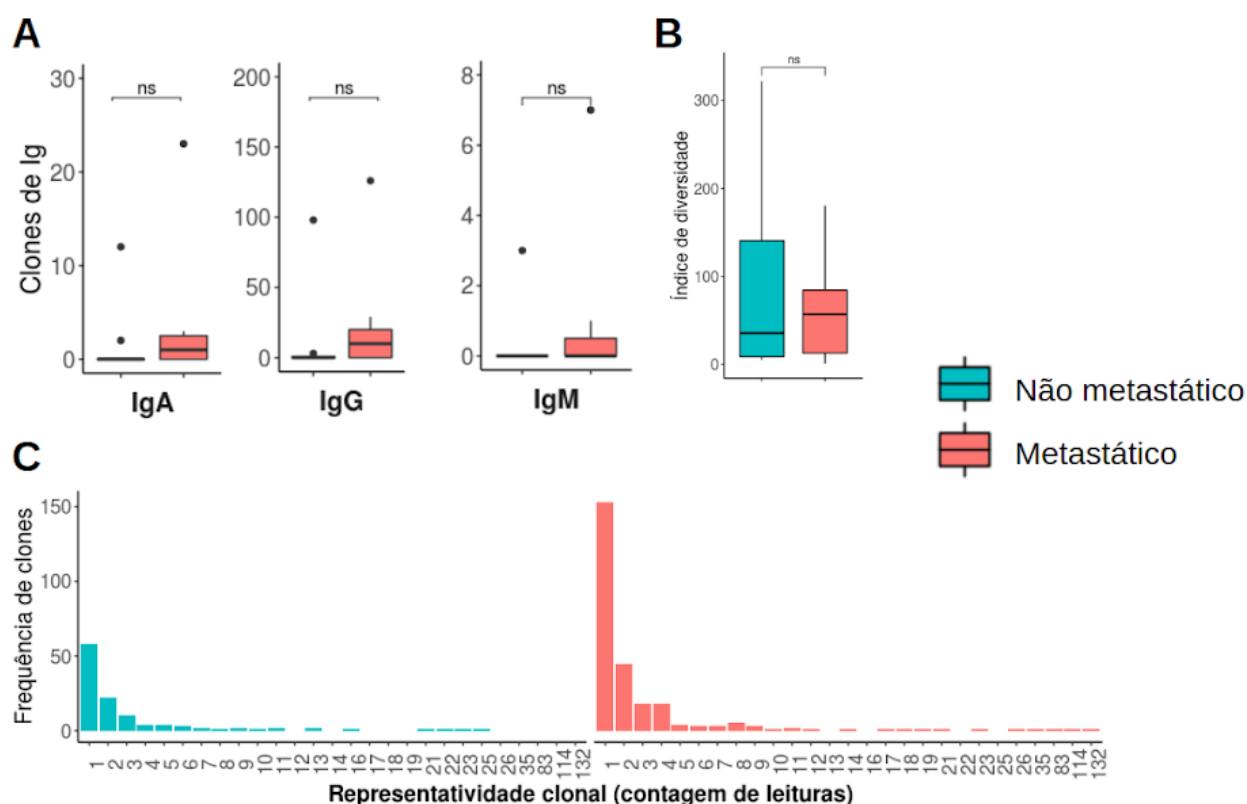


Figura 4.13 Repertório de BCR entre os grupos. **A.** Número de clones de imunoglobulina A (IgA), G (IgG) e M (IgM) identificadas entre os grupos. Não houve diferença significativa entre os grupos no teste de comparação de médias pelo método Wilcoxon (Mann-Whitney). **B.** Cálculo do índice de diversidade *inverse Simpson* não evidenciou diferença entre os grupos. **C.** Frequência da representatividade clonal. Cada clone possui um número de leituras (*reads*) que suporta sua identificação e serve como uma medida indireta da expansão desse clone. Foi excluída dos gráficos A

e C uma amostra *outlier* que contava com 701 clones de IgG, 177 de IgA e 3 de IgM, mas seus valores foram contabilizados para o teste de médias.

4.6 Perfil mutacional e de neoantígenos

Dentre os oito genes mais frequentemente mutados em tumores de cabeça e pescoço com impacto na tumorigênese, *TP53*, *FAT1*, *CDKN2A*, *PIK3CA*, *NOTCH1*, *KMT2D*, *NSD1* e *CASP8* (CANCER GENOME ATLAS NETWORK, 2015), 7 dos 8 foram encontrado mutados nos tumores não metastáticos e 3 dos 8 no grupo metastático (Fig 4.14). Em algumas amostras foram identificadas mais de uma mutação no mesmo gene (*TP53*, *FAT1*, *KMT2D* e *NSD1*) e 2 das 22 amostras não possuíam nenhum desses oito genes mutados (Fig 4.14 - NE).

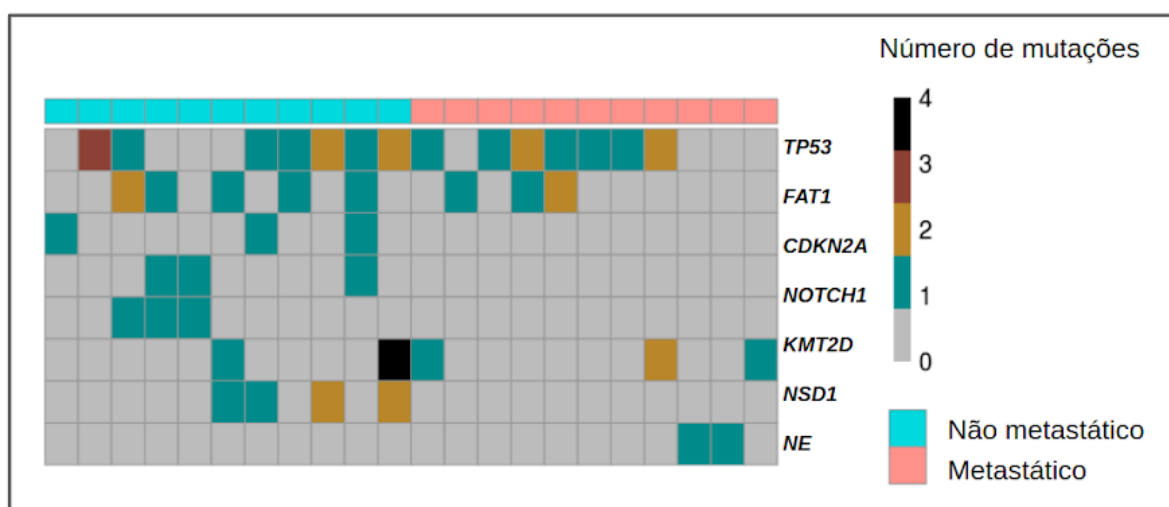


Figura 4.14 Frequência de mutações nas coortes analisadas para cada um dos oito genes mais frequentemente mutados em tumores de cabeça e pescoço. Os genes estão em ordem decrescente de frequência. NE. Não encontrada.

A carga de mutação exônica, intrônica e total não apresentou diferença estatística entre os grupos (comparação de médias pelo método de Wilcoxon, Figura 4.15-A) e nem na comparação de NeoAg e NeoAg derivados de TAA (comparação de médias pelo método Wilcoxon, Figura 4.15-B). Apesar disso, foi observada uma tendência de maior carga total de mutação (intrônica e exônica) no grupo não metastático (mínimo de 16, máximo de 1.534 e mediana de 118) em relação ao grupo metastático (mínimo de 8, máximo de 316 e mediana de 54). Os NeoAg derivados de mutação seguiram a mesma tendência entre o grupo não metastático (mínimo de 23, máximo de 491 e mediana de 51) e metastático (mínimo de 5, máximo de 133 e mediana de 29). Já os NeoAg derivados de proteínas de

expressão aberrante (TAA) seguiram uma tendência inversa entre o grupo não metastático (mínimo de 4, máximo de 48 e mediana de 23) e metastático (mínimo de 0, máximo de 88 e mediana de 32).

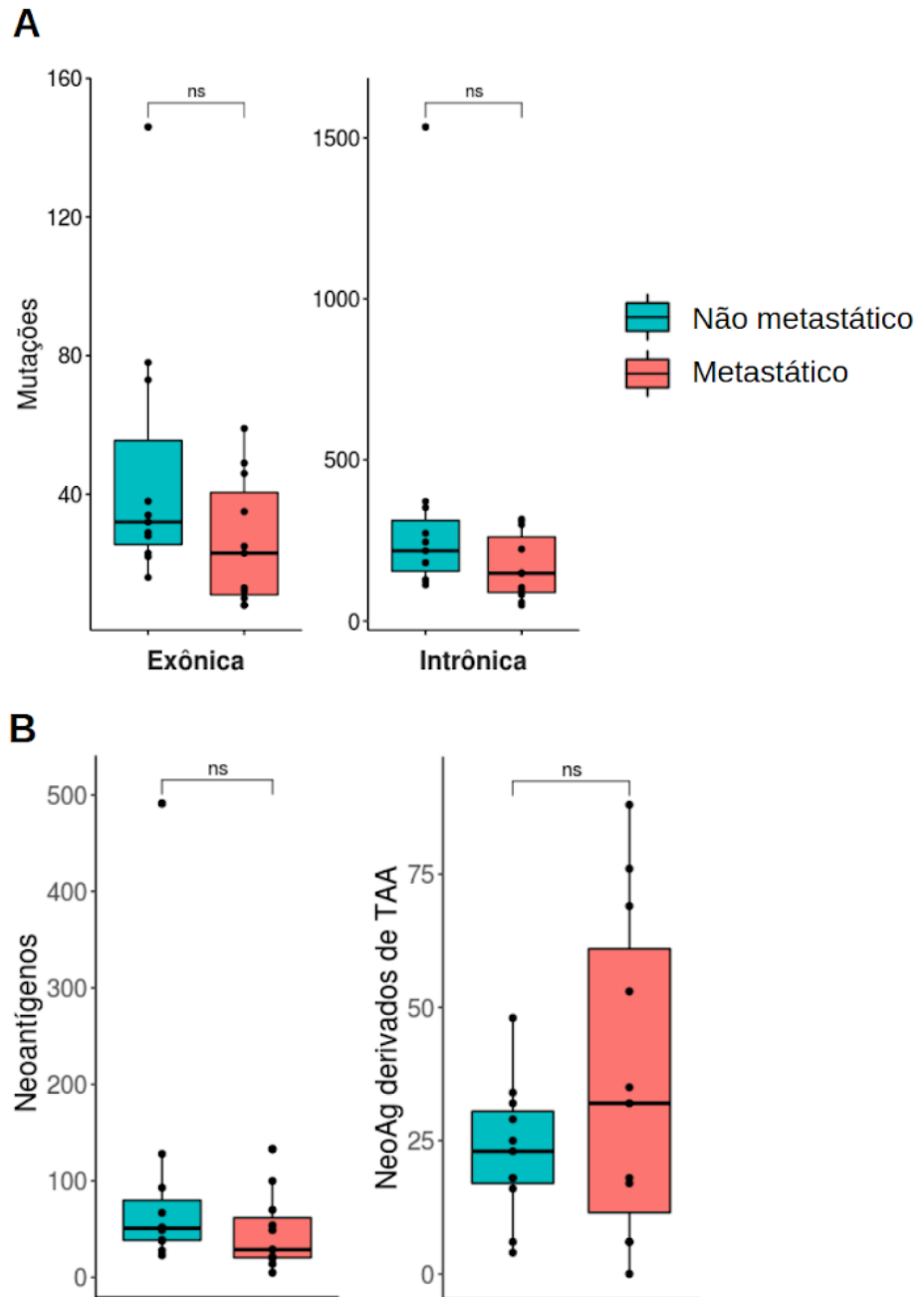


Figura 4.15 Carga de mutação e NeoAg da coorte do TCGA. **A.** Número de mutações exônicas (esquerda) e intrônicas (direita). **B.** Número de NeoAg derivados de mutação (esquerda) e derivados de TAA - antígenos de expressão aberrante (direita).

A carga de mutação exônica correlaciona positivamente com a carga de NeoAg ($\rho > 0,85$ e valor de $p < 0,01$, Spearman; dado não mostrado). A Figura 4.16-A está ordenada de forma crescente em carga de NeoAg e mostra essa relação positiva entre carga de mutação e NeoAg. Três amostras do grupo não metastático e

uma do metastático possuem um ou mais alelo de HLA em homozigose (Fig 4.16-A; setas). Foi encontrada uma mutação não sinônima em *PSME2*, uma proteína do proteassomo, e em *ERAP1*, da via de processamento, na amostra indicada por um asterisco (*). Não foram observadas diferenças significativas na expressão de *TAP1*, *TAPBP* (tapasina) e *B2M* entre os grupos. Na maioria das amostras, as mutações encontradas são, na maior parte, não sinônimas (~60%). Contudo, em três casos do grupo metastático essa proporção chega a < 40% (Fig 4.16-B). A assinatura mutacional das amostras se mostrou relativamente consistente entre os grupos e amostras, com prevalência de conversões Citosina para Timina (Fig 4.16-C). Foram encontradas mutações não sinônimas em genes de reparo ao dano ao DNA em 4 amostras, exatamente aquelas que possuem maior carga de NeoAg (Fig 4.16-D).

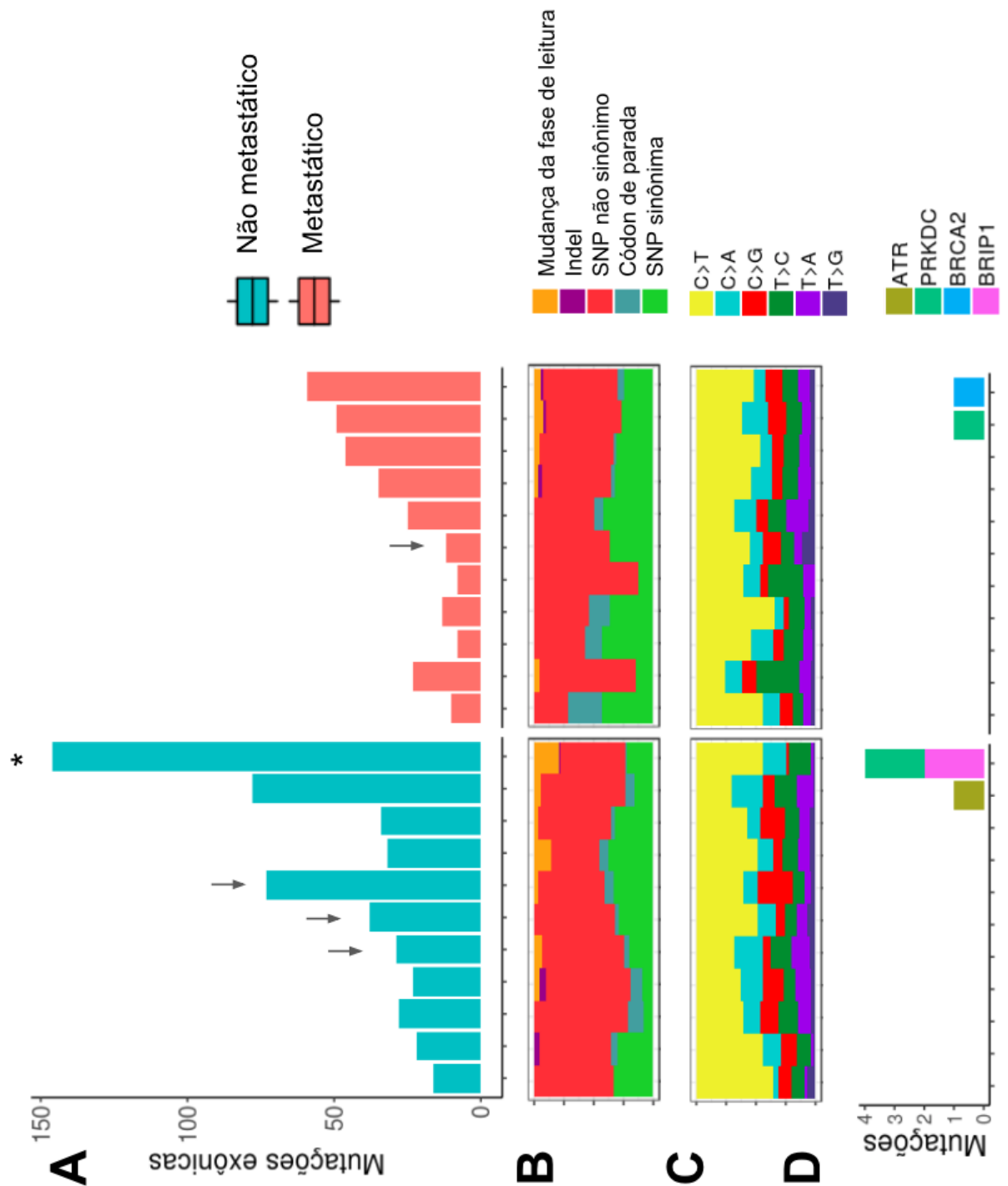


Figura 4.16 Perfil mutacional da coorte TCGA. **A.** Número de mutações exônicas por amostra ordenada pela carga de NeoAg; as setas indicam aquelas com homozigose em pelo menos um alelo de HLA; o asterisco indica mutação no gene da via de processamento e apresentação de antígeno, *PSME2*. **B.** Proporção relativa do impacto das mutações; não foram encontradas mudanças no quadro de leitura. **C.** Perfil de transversão (Citosina>Adenina, Citosina>Guanina; Timina>Adenina; Timina>Guanina) e transição (Citosina>Timina; Timina>Citosina) das amostras com a proporção relativa de cada alteração. **D.** Mutações em enzimas de reparo ao dano de DNA.

4.7 Escape tumoral

A partir dos dados gerados até então, foi realizada uma análise holística combinando informações de carga de NeoAg, expressão de moléculas inibitórias e estimulatórias, proteínas da via de apresentação de antígeno, citocinas, sub-populações celulares imuno-regulatórias e número de clones de cadeias CDR3 do TCR. Os genes selecionados foram agrupados intragrupo utilizando distância euclidiana. Não houve diferença estatística entre os genes analisados entre os grupos pelo teste de Wilcoxon (Mann-Whitney).

A amostra com mutação na via de processamento e apresentação de antígeno e que possui também a maior carga predita de NeoAg (2^a amostra da esquerda para a direita no grupo não metastático) apresentou a menor expressão de *B2M* do grupo (Fig 4.17-A). Ela é uma das duas mais reguladas negativamente para esses marcadores e que possuem um dos menores número de clones de CDR3 identificados para o grupo (Fig 4.17-C, número 1). Cinco amostras desse grupo estão reguladas positivamente para moléculas inibitórias e estimulatórias (Fig 4.17-A). Essas amostras possuem o maior número de clones de CDR3 identificados (Fig 4.17-C, grupo número 2).

No grupo metastático, uma amostra com regulação negativa da maioria dos receptores estimulatórios e inibitórios (Fig 4.17-B) não teve clone de CDR3 identificado (Fig 4.17-C, número 3). Do mesmo modo que no grupo não metastático, três amostras com maior expressão de moléculas inibitórias e estimulatórias dentro do grupo (Fig 4.17-B, número 4) foram as que tiveram os maiores números de clones de CDR3 identificados (Fig 4.17-C, número 4).

A expressão de *TAP1* acompanhou a de *TAP2* na grande maioria dos casos, ao contrário da expressão de *B2M*. Nos *cluster* com maior regulação positiva (2 e 4), houve também uma assinatura mais forte de Treg e MDSC. Não houve agrupamento aparente por carga de NeoAg, consumo de álcool e sítio anatômico.

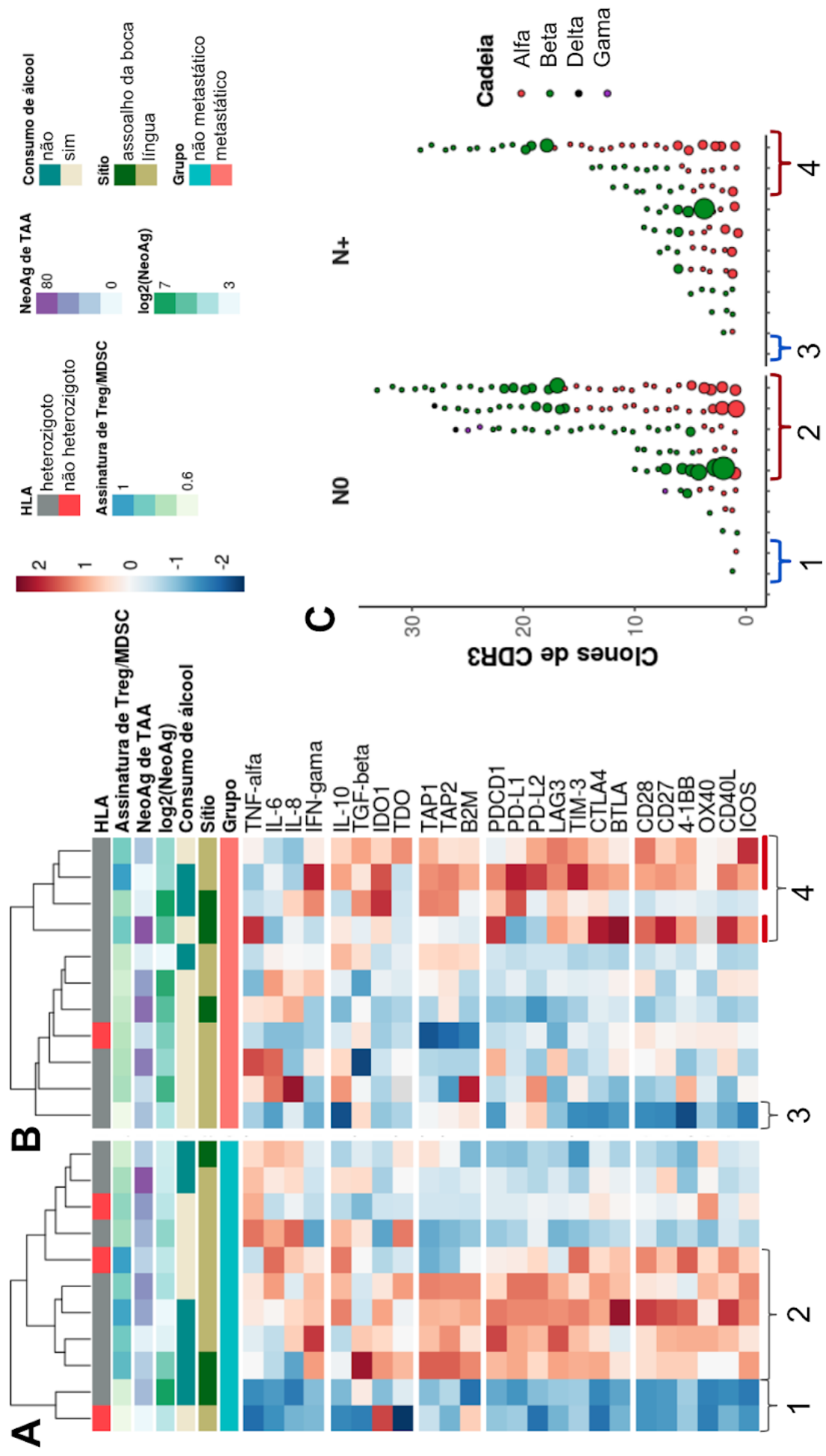


Figura 4.17. Panorama imunológico da coorte TCGA. Mapa de calor do grupo não metastático **(A)** e metastático **(B)** mostrando a expressão de genes que codificam citocinas pró-inflamatórias, fatores solúveis anti-inflamatórios, proteínas da via de apresentação de antígeno, moléculas inibitórias e estimulatórias. Os painéis superiores indicam se o indivíduo possui ou não ao menos um alelo de HLA em heterozigose, a assinatura de células supressoras (MDSC e Treg) de Charoetong e colaboradores (CHAROENTONG et al., 2017), carga de NeoAg derivado de TAA, carga de NeoAg em log2, histórico de consumo de álcool, sítio anatômico do tumor e grupo ao qual pertence. **C.** Rearranjos da região CDR3 do TCR contendo número de clones (eixo y), cadeia do receptor (cores) e representatividade clonal. Treg. Célula T reguladora. HLA - Antígeno leucocitário humano. MDSC - Célula supressora de origem mieloide. NeoAg - Neoantígeno. TAA - Antígeno associado ao tumor. CDR3 - Região determinante da complementariedade 3.

5 Discussão

A coorte de validação foi selecionada a partir das amostras do TCGA de modo a ficar o mais semelhante possível àquela do ACC. Contudo, não havia tumores metastáticos em estadiamento t1 suficientes para a equiparação. O estadiamento patológico dos tumores que na coorte ACC consiste inteiramente de tumores t1 e a coorte TCGA é majoritariamente composta por tumores em t2. A fim de excluir o viés introduzido por amostras provenientes de indivíduos do sexo feminino, três amostras foram excluídas de um dos grupos da coorte, restando somente indivíduos do sexo masculino em ambas. O sítio anatômico do grupo metastático da coorte ACC está igualmente dividido em língua e assoalho da boca, ao contrário do grupo metastático da coorte do TCGA que possui a maior parte na língua (Tabelas 3.2 e 4.1). Uma abordagem alternativa seria pesquisar os genes diferencialmente expressos entre o mesmo sítio anatômico. Diferenças intrínsecas das coortes, sítios anatômicos, sexo e grau do tumor influenciam no resultado. Além disso, as amostras da coorte do TCGA não provêm de indivíduos brasileiros e somente 1 dos 22 indivíduos é de origem latina ou hispânica declarada. O *background* genético acumulado pelas etnias interfere no desenvolvimento de tumores e isso pode ajudar a justificar o baixo número de DEG em comum encontrados. (UPPSALA UNIVERSITET, 2015).

Os problemas durante o sequenciamento da corrida 2 diminuíram a confiança na chamada das bases, o que é observado pelos valores mais baixos na escala Phred de qualidade (Fig 4.1-A e C). Para a identificação de mutações, o consenso é aceitar bases com um valor phred de qualidade igual ou superior a 20, que é correspondente à faixa amarela na Figura 4.1 (GATK, 2014). Devido à baixa qualidade do sequenciamento da plataforma SOLiD, não foi realizada chamada de variante para a coorte ACC. O final das sequências de ambas as corridas possui qualidade menor que o restante das leituras; uma tendência de algumas técnicas de sequenciamento, pois a química da reação não possui um rendimento prático de 100% e uma minoria das leituras de um *cluster* perdem sincronia com as demais (ECSEQ BIOINFORMATICS, 2017). Por conta da parada precoce em uma das corridas do sequenciamento, não foram utilizados parâmetros mais estritos de filtragem das sequências. Filtros menos permissivos aumentariam a qualidade e

confiança das sequências obtidas, mas diminuem drasticamente o número de sequências habilitadas para análises subsequentes. O alinhamento de sequências em um genoma de referência para o cálculo da expressão gênica é influenciado pelo número, comprimento e qualidade das sequências. Desse modo, é de mais valia preservar um número mais elevado de leituras, mesmo que com qualidade um pouco inferior, do que utilizar filtros muito restritivos. O mesmo não é verdade para uma análise de chamada de variantes, na qual a certeza na chamada de uma base é importante para se determinar mutações (LI; DEWEY, 2011) (REF pro .vcf). Por se tratar de um dado oriundo da plataforma SOLiD, as leituras possuem qualidade inferior às geradas pela plataforma Illumina. Isso não permite que o dado seja analisado pelos programas e pacotes mais recentes da mesma forma. Para poder recorrer a estas análises, foi realizada a conversão dos dados para o formato .fasta, que não contém indicadores de qualidade de base, somente a sequência (Fig 3.1). Trabalhos recentes demonstraram que o fato das bases não possuírem qualidade não interfere significativamente na qualidade do alinhamento (LI; DEWEY, 2011).

O efeito de lote observado entre as corridas pode interferir na interpretação dos resultados se não for adequadamente considerado e abordado metodologicamente. O fato da coorte do ACC possuir o mesmo número de tumores metastáticos em ambas as corridas auxilia a amenizar esse efeito, já que é o fenótipo pesquisado. Durante o cálculo dos genes diferencialmente expressos foi necessária a inclusão das corridas como parâmetros de covariância, assim como a metastização. Apesar disso, o efeito de lote não é removido das amostras, ele é apenas considerado durante o cálculo de expressão diferencial e por isso ele continua presente como ilustrado pela Figura 4.3 (LOVE; ANDERS; HUBER, 2019).

A sobreposição dos genes diferencialmente expressos (DEG) entre as duas coortes foi pequena (Fig 4.5). Contudo, algumas diferenças marcantes nos aspectos clínicos das coortes podem ter colaborado para essas diferenças como discutido no tópico anterior. Apesar da pequena sobreposição de DEG, o dado fornecido pelas vias enriquecidas é capaz de suprir essa limitação, já que utiliza 6.000 genes (3.000 com maior e 3.000 com menor *log2fold change*). Com exceção do gene *ADH1B* na coorte ACC, não houve diferença entre os dois genes identificados como DEG pelo DESeq2 no teste de comparação de médias (Fig 4.6). Isso pode ser devido à modelagem que o DESeq2 utiliza para normalizar a dispersão da contagem bruta

dos genes (Fig 4.4) que reduz a dispersão dos genes conforme o nível de expressão para identificar DEG com maior precisão e minimizar ruídos em coortes pequenas.

A álcool desidrogenase 1B (*ADH1B*) é uma enzima que oxida o etanol, retinol e outros substratos alcoólicos no organismo. O consumo de etanol é considerado pelo IARC como um agente carcinogênico, especialmente em órgãos que entram em contato direto com ele ou o metabolizam, como a cavidade oral, fígado e o trato digestivo superior (BAAN et al., 2007). A maior expressão de *ADH1B* no grupo metastático pode estar associada ao maior consumo de álcool. Dos pacientes do TCGA sem histórico de consumo de etanol, 6 (total de 9) não tiveram desfecho de metástase e dos que declararam consumo de etanol 8 (total de 13) tiveram desfecho de metástase. Os dados clínicos referentes ao consumo de álcool da coorte ACC ainda não foram analisados. Variantes alélicas da *ADH1B* possuem atividades metabólicas diferentes e a identificação dos mesmos pode ser valiosa para elucidar o envolvimento desse gene no fenótipo metastático (POLIMANTI; GELERNTER, 2018).

O gene *PIWIL2* (Piwi Like RNA-Mediated Gene Silencing 2) pertence à família argonauta, que contém genes envolvidos no silenciamento de RNA provenientes de elementos transponíveis e são responsáveis pela manutenção de características de células germinativas por conta da participação na modulação epigenética, remodelamento da cromatina e regulação do ciclo celular (SASAKI et al., 2003; YE et al., 2010). Alguns trabalhos descrevem a relação desse gene com a carcinogênese em diferentes tumores (LITWIN et al., 2018; YE et al., 2010), e inclusive demonstram que seu bloqueio por anticorpo inibe o crescimento tumoral (GAO et al., 2016). A expressão de *PIWIL2* é basicamente restrita ao testículo (HUMAN PROTEIN ATLAS, [s.d.]), o que o torna um antígeno imunogênico quando expresso em outros tecidos. Considerando os dois grupos da coorte, a expressão de *PIWIL2* é maior no grupo não metastático que possui tumores classificados no estágio t3 ou superior enquanto o grupo com tumores t2 ou inferior possui expressão mais baixa. Tendo em vista o papel imunogênico de *PIWIL2*, tumores com menor expressão desse gene poderiam ter mais facilidade para evadir da vigilância imunológica e, conseqüentemente, formar metástases. Um fato contraditório com o desfecho não metastático é a recente associação da expressão de *PIWIL2* e

proteínas da família *MAGEA* com a regulação positiva de genes relacionados à transição epitélio-mesênquima e metástase (YANG et al., 2015).

O padrão do agrupamento dos DEG indica uma homogeneidade das amostras na coorte ACC, pois a primeira separação do dendograma agrupa 7 amostras do grupo metastático (Fig 4.7-A). As separações seguintes agrupam as amostras em 3 clusters de uma mesma condição, mesmo que 3 amostras da condição metastática estejam entre amostras do outro grupo. Essas 3 amostras do grupo metastático são compostas pelas 2 únicas pacientes do sexo feminino e outra do sexo masculino. As 3 amostras tem origem do mesmo sítio anatômico (língua - C02) e nenhuma outra particularidade clínica associada identificada. As amostras do assoalho da boca estão dentro do mesmo *cluster*, o que reforça o efeito que o sítio anatômico pode ter sobre os DEG, ainda mais se considerado o efeito de lote (corrida 1).

A coorte do TCGA também parece homogênea, pois 9 amostras do grupo não metastático estão agrupadas após a primeira separação do dendograma (Fig 4.7-B). No outro lado do dendograma, um *cluster* contendo 7 amostras do grupo metastático fica evidenciado, separado do demais. Isso indica que essas amostras são similares entre si no padrão de expressão gênica, ao contrário das 4 outras amostras desse mesmo grupo. A diferença nos genes regulados positivamente pelo grupo metastático não é homogênea entre a coorte do TCGA e contém duas amostras *outliers* (Fig 4.7-B, 10^a e 11^a amostra da esquerda para a direita). O nível de expressão destoante dessas duas amostras afeta a escala de cores para os outros genes desse grupo. Uma das duas amostras do grupo não metastático entre as amostras do outro grupo é a única que possui etnia latina (Fig 4.7-B, rótulo verde). As amostras oriundas do assoalho da boca não ficaram agrupadas, o que sugere que a distribuição de amostras deste sítio entre os grupos pode ter minimizado esse efeito e que o agrupamento observado na coorte ACC pode ter sido devido ao efeito de lote.

O volcano plot da coorte ACC possui um conjunto de genes da família *MAGEA* (*MAGEA12*, *MAGEA2*, *MAGEA2B*) regulado positivamente no grupo não metastático (Fig 4.8-A). Esses genes estão localizados no mesmo *locus* e são regulados em *cluster* por mecanismos de metilação, o que pode explicar sua expressão conjunta (DE SMET et al., 1999). Os genes da família *MAGEA* codificam

proteínas que pertencem a uma categoria de antígenos de expressão restrita (tecido ou embriogênese-específica), também conhecidos como antígenos associados a tumor (TAA). A maior expressão desses genes no grupo não metastático pode ser um fator que colabora para a ausência de metástase, ao passo que o SI é responsivo a esses antígenos (ZAJAC et al., 2017). Além disso, já foram associados à desregulação de p53 e maior potencial proliferativo, aspecto característico do grupo não metastático (ZAJAC et al., 2017). A Interleucina 1 alfa (IL1A), uma citocina pró-inflamatória, também foi encontrada como uma das mais reguladas positivamente no grupo não metastático. Ao contrário da IL1 beta, a IL-1 alfa tem localização no citosol ou membrana celular (VORONOV et al., 2003); a IL-1 beta promove a invasividade tumoral, enquanto a IL-1 alfa não (APTE et al., 2000). Sua relação com a promoção da inflamação pode favorecer a geração de CTL e o controle das células tumorais de modo a restringir a localização do tumor, mesmo que não o elimine por completo (APTE et al., 1993). Ainda na coorte ACC, um dos genes mais regulados positivamente no grupo metastático foi o Fator de crescimento de fibroblasto (*FGF*). Esse gene tem associação com aumento de motilidade e invasividade em diferentes tumores e um membro dessa família, o *FGF-2*, tem atuação sinérgica com o VEGF (KORC; FRIESEL, 2009). A maior expressão desse gene pelo grupo metastático evidencia o fenótipo do grupo por conta da participação do *FGF* e dos CAF na promoção da angiogênese e invasividade.

A coorte do TCGA possui dois genes mais expressos no grupo metastático, a alfa 1 e alfa 2 Actina (*ACTA1* e *ACTN2*). A actina se organiza na célula em microfilamentos de actina, um componente do citoesqueleto celular que se localiza próximo à membrana celular e que auxilia na migração da célula. Essa função de migração celular está relacionada à invasividade de tumores e geração de metástases (YAMAGUCHI; CONDEELIS, 2007). Em tumores de cavidade oral, a reorganização dos microfilamentos de actina é relatada como consequência da regulação de genes associados a esse fenótipo, como a Galectina-1 (SHARMA et al., 2013). O aumento da reorganização do citoesqueleto celular facilita a migração e invasividade celular e ajuda a explicar o desfecho metastático. Já no grupo não metastático da coorte TCGA, um dos genes mais expressos foi o CD300E, uma glicoproteína de superfície com domínios de imunoglobulina expressa em células mieloides, principalmente em monócitos. O CD300E foi descrito como um receptor

de ativação de respostas imunológicas e, quando em interação com um ligante extracelular, transduz sinais de ativação que regulam positivamente co-ativadores (e.g. CD25) e citocinas pró-inflamatórias (e.g. TNF-alfa) (BRCKALO et al., 2010). Sua expressão pode ser proveniente do infiltrado imunológico, mas também do próprio tumor, já que não é possível distinguir a origem celular da expressão no tipo de dado analisado.

As vias de queratinização e formação do envelope cornificado (estrato córneo), enriquecidas no grupo não metastático, são as únicas em comum entre as coortes. Elas podem refletir a maior integridade dos tecidos epiteliais do dorso da língua por conta do menor contato com o álcool (GANEM-QUINTANAR et al., 1998) e reforçar a maior expressão de *ADH1B* no grupo metastático. Essa possível associação do desfecho de metástase com o consumo de álcool será explorado em trabalhos futuros. As vias enriquecidas no grupo metastático são majoritariamente formadas por assinaturas relacionadas ao músculo cardíaco e esquelético da coorte TCGA (Fig 4.9, superior). Essa relação pode ser decorrente da expressão de DEG como a *ACTA1*, *ACTN2* e outros relacionados ao citoesqueleto, e não necessariamente refletir um fenótipo de cardiomiócito.

Vias relacionadas ao SI estão enriquecidas no grupo não metastático, como a sinalização de interferon alfa e beta, interações imunoregulatórias, degranulação de neutrófilos, processamento e apresentação de antígenos e citotoxicidade mediada por célula NK (Figura 4.9, inferior). Os interferon alfa e beta são produzidos por diversas células e possuem atividade antitumoral. A atividade desta via pode indicar uma atividade antitumoral maior no grupo não metastático responsável pela restrição do tumor ao sítio inicial. Contudo, já foi descrito que a sinalização por esses interferons favorece perfis migratórios, o que contradiz com o desfecho não metastático (PARKER; RAUTELA; HERTZOG, 2016). A degranulação de neutrófilos sugere uma atividade citotóxica que os mesmos estão efetuando de modo mais pronunciado no grupo não metastático e, assim como o descrito anteriormente para o interferon, auxilia no controle da evasão tumoral (ROSALES, 2018). Do mesmo modo, a citotoxicidade mediada por células NK pode favorecer a eliminação e controle de células tumorais. Neste caso, esta atividade não parece ter sido capaz de impedir o surgimento e crescimento do tumor. Essas respostas do SI em curso não são dependentes da apresentação de antígenos, outra via mais regulada

positivamente no grupo não metastático. O menor processamento e apresentação de antígenos dificulta a resposta mediada por células T e favorece o escape tumoral. Logo, a menor atividade desta via no grupo metastático pode auxiliar o escape tumoral e geração de metástase no início do desenvolvimento do tumor.

Duas vias enriquecidas no grupo não metastático dizem respeito à biogênese ribossomal e assinatura de ribossomo. A síntese de proteínas ribossomais é um processo energeticamente caro, principalmente se considerado células em contínuo crescimento, e alterações na biogênese ribossomal foram relacionadas ao desfavorecimento da progressão tumoral via estabilização de p53 (PELLETIER; THOMAS; VOLAREVIĆ, 2018). Como o grupo não metastático possui tumores maiores, pode-se esperar que a via de biogênese ribossomal esteja aumentada, já que será útil à formação de novas proteínas. Por fim, é necessária uma avaliação cuidadosa das proteínas ribossomais diferencialmente expressas em cada grupo já que diversas destas proteínas estão associadas positivamente e negativamente com a metástase (DE LAS HERAS-RUBIO et al., 2014).

As assinaturas de populações imunes identificadas pelo xCell foram estabelecidas a partir de dados de transcriptomas humanos e fornecem o enriquecimento para uma dada população celular, não a fração em que essa população está presente na amostra (ARAN; HU; BUTTE, 2017). Dentre as 64 assinaturas disponíveis, foram mostradas as quantificações daquelas com pelo menos metade das amostras possuindo valor não nulo de enriquecimento (Fig 4.10).

A coorte ACC não apresentou uma separação boa dos grupos com assinaturas mais características de um grupo ou outro. Contudo, algumas assinaturas parecem enriquecidas no grupo não metastático, como a de células epiteliais, queratinócitos e sebócitos (Fig 4.10-A). A assinatura de queratinócitos, sebócitos e células epiteliais pode refletir as características epiteliais e glandulares da superfície da língua, que é composta por epitélio estratificado (SOUTHERN ILLINOIS UNIVERSITY, 2012), e é o único sítio anatômico que compõe esse grupo (vide legenda). Em contraste, algumas amostras do grupo metastático possuem assinaturas de células endoteliais, endoteliais microvasculares e célula-tronco hematopoiética (Fig 4.10-A). Essas populações celulares podem ter relação com a

angiogênese que, por sua vez, favoreceu o perfil invasivo do grupo metastático. A presença de assinatura para célula-tronco hematopoiética não é esperada em tecidos extra-medulares, mas pode estar refletindo outras causas, como será discutido adiante. Houve uma separação grande das amostras devido ao efeito de lote, como observado na legenda da corrida, o que atrapalha a interpretação do resultado.

A coorte TCGA apresentou uma separação melhor entre os grupos, com um pequeno *cluster* de amostras com metástase contendo enriquecimento de assinaturas para populações como células endoteliais, neurônios, miócitos, músculo esquelético e célula-tronco hematopoiética (Fig 4.11-B). Do mesmo modo que na coorte ACC, as assinaturas de células endoteliais estão preferencialmente presentes nas amostras do grupo metastático e pode refletir um processo angiogênico. A língua é composta basicamente de músculo, mas também por diversos receptores sensitivos, o que pode explicar as assinaturas de células musculares e neuronais, vide legenda. Esse resultado corrobora com o encontrado nas vias enriquecidas, no qual diversas vias celulares de contração muscular foram identificadas (Fig 4.9). A população de célula-tronco hematopoiética, assim como na coorte ACC, está mais enriquecida no grupo metastático e, como veremos adiante, possui correlação com o grupo metastático. É importante notar a separação de um grupo de 4 amostras sem consumo de álcool que possuem uma assinatura modesta de células epiteliais, sebócitos e queratinócitos. Elas podem refletir a integridade do tecido epitelial da língua por conta da ausência de consumo de álcool.

De modo similar às quantificações, estão apresentadas as correlações entre populações e desfechos aquelas com módulo de correlação de Pearson $\geq 0,4$ para filtrar possíveis ruídos (Fig 4.11). Algumas populações presentes nas análises de correlação não estiveram nas quantificações de assinaturas (Fig 4.10), devido aos critérios de filtro. A única população presente em ambas as coortes foi o linfócito T CD4 de memória que, inclusive, possui uma das maiores correlações negativas com o desfecho. Um dos principais componentes dessa população é o linfócito Th que dá suporte à atividade de células efetoras (e.g. CTL) na eliminação de células (perfil Th1) e na produção de anticorpos (Th2) (KENNEDY; CELIS, 2008). O enriquecimento dessa população na condição não metastática pode ter relação com seu papel antitumoral no sentido de evitar a evasão tumoral, mas que não se

mostrou tão efetivo no controle do crescimento tumoral, já que os tumores não metastáticos são maiores. Os mecanismos de eliminação de células tumorais podem ter sido mais efetivos na condição metastática, pois impediram o maior crescimento tumoral, mas por outro lado podem ter sido mais favoráveis à metastização. Existem evidências da participação de células do microambiente no favorecimento da metástase, mas pouco se sabe sobre o papel da intensidade da resposta antitumoral na metastização (WANG et al., 2017). Os dois subtipos de linfócito T CD4 de memória, memória central e efetora, tiveram correlações muito baixas com o desfecho em ambas as coortes. Eles poderiam dar mais informações sobre as características predominantes de resposta da população CD4 de memória.

Três subpopulações de linfócitos T CD8 estão associadas, apesar de fracamente, com a ausência de metástase linfonodal em ambas as coortes. Os linfócitos T CD8 são os principais efetores na resposta antitumoral celular e os resultados mostram indícios de que estejam associados a um desfecho livre de metástase linfonodal. Apesar desse resultado estar em conformidade com o enriquecimento da população T CD4 de memória na condição não metastática, pode parecer um contrassenso o enriquecimento dessas populações não ser capaz de impedir o desenvolvimento tumoral exacerbado, pelo contrário, uma vez que essas assinaturas estão presentes em tumores de maior volume. Isso pode ser devido ao próprio tamanho tumoral que desencadeia mais processos de necrose, inflamação e recruta mais células inflamatórias ou também devido à maior atividade de vias oncogênicas, como visto na Figura 4.10.

As outras assinaturas encontradas não estão presentes em ambas as coortes ou não possuem uma correlação tão significativa. Exceção à essa regra é a assinatura de célula-tronco hematopoiética (HSC) que possui associação de 0,6 com a metástase. Não é esperado encontrar essa população celular em tecidos extra-medulares sob condições fisiológicas normais e em indivíduos adultos. Contudo, em casos de doenças hematopoiéticas, infecções, processos de inflamação sistêmica e neoplasias, a hematopoiese extra-medular pode ocorrer em diferentes tecidos, tal como o linfonodo (ELMORE, 2006). É possível que essa assinatura no tumor primário esteja relacionada com uma hematopoiese extramedular no linfonodo sentinela. Outra possibilidade é que a assinatura de HSC reflita, na verdade, uma assinatura de célula-tronco tumoral. Apesar da assinatura para HSC possuir

marcadores característicos da linhagem hematopoiética (e.g. CD34), os marcadores são avaliados como um todo e não consideram as células-tronco tumorais. Importante notar que a correlação de Pearson para essa população (Fig 4.11) pode estar sendo superestimada pela amostra *outlier* muito enriquecida para essa assinatura (Fig 4.10).

Outra população que apareceu isolada, mas com correlação de 0,77 com o desfecho metastático e dessa vez na coorte TCGA, foi o Pró-linfócito B (Fig 4.11-B). Um resultado similar de população B (Linfócito B de memória) foi encontrado nesta coorte usando outro algoritmo e reforça esse achado. Existem relatos em outros tumores da participação de linfócitos B no aumento do fluxo do linfonodo drenante e, conseqüente, da metástase; além de fatores secretados por eles que favorecem a transição epitélio mesênquima (KOIZUMI et al., 2013; RUDDELL et al., 2011). A assinatura de músculo liso correlacionando fracamente com o grupo não metastático pode indicar a presença de músculo liso vascular e irrigação sanguínea (Fig 4.11-A). A maior assinatura de células endoteliais microvasculares no grupo metastático (Fig 4.10) está em conformidade com o fato de tumores maiores, como os do grupo, demandem maior aporte de O₂.

Esses fatores reforçam a hipótese de que o sítio anatômico, língua e assoalho da boca, podem ter assinaturas de DEG e população do microambiente tumoral bem diferentes. Além disso, o consumo de álcool é um fator característico das amostras com metástase linfonodal, como já citado, e sua influência nas assinaturas de populações celulares e DEG merece um estudo individual.

Os agrupamentos de BCR oriundos de um mesmo clone são menos fielmente identificados, pois o linfócito B ativado sofre hipermutação somática. Isso faz com que exista uma variedade de BCR similares que não são agrupados pelo programa no mesmo clone. Outro fator complicador da hipermutação somática é o ruído que ela causa durante a fase de alinhamento, já que os emparelhamentos errados permitidos pelo programa durante o alinhamento podem levar a uma estimativa exagerada do número de clones (BOLOTIN et al., 2015). Uma limitação do método é a ausência de informação acerca do subtipo de linfócito T (CD4 e CD8) a qual o receptor identificado pertence. Alguns dos fatores que influenciam na descoberta de clones de TCR e BCR são: o tamanho da leitura (*read*); a quantidade de infiltrado imunológico; e a cobertura do sequenciamento. O número de leituras variou de 92

milhões a 197 milhões entre as amostras e não mostrou relação aparente com o número de clones encontrados (dado não mostrado). O tamanho de leitura (*read*) recomendado para identificação de clones de TCR e BCR é de 150bp, mas o das amostras do TCGA é de 50bp. Contudo, o tamanho das leituras não é tão relevante na detecção de clones de TCR como de BCR, por conta da hipermutação somática.

Apesar do número de clones provenientes de cadeias alfa, beta e da diversidade dos receptores não apresentar diferença entre os grupos, o grupo não metastático apresenta maior número de clones mais representados (Fig 4.12). Essa representação clonal se refere ao número de leituras (*reads*) que suportam a existência desse clone e pode ser entendida como uma medida indireta da expansão do mesmo. Clones expandidos podem estar respondendo a um antígeno tumoral dominante no microambiente. O fato do grupo não metastático possuir um número superior de clones de TCR superrepresentados sugere uma relação desse desfecho com uma resposta de linfócito T expandida (Fig 4.13-C).

Não foi possível identificar diferença entre o número de clones de BCR e diversidade dos mesmos (Fig 4.13-A-B). Uma das amostras do grupo não metastático teve um número de clones identificados que totaliza 70% do total de clones de BCR de todas as amostras dos dois grupos. Esse número exacerbado de clones pode ser devido ao viés introduzido pela hipermutação somática ou mesmo por conta de alguma estrutura linfoide terciária que foi coletada junto da amostra tumoral. Considerando a possibilidade de estrutura linfoide, foi verificada a lâmina da biópsia dessa amostra depositada no TCGA e notou-se uma estrutura sugestiva de órgão linfoide terciário (Anexo 3). Outras três amostras do TCGA possuem estruturas sugestivas de órgão linfoide terciário, uma do assoalho da boca e duas da língua. Ao contrário do que foi identificado com os TCR, a maior representatividade clonal no grupo metastático pode ter relação com uma maior resposta B em curso (Fig 4.13-C). Wirsing *et al.* reportaram a identificação de estruturas linfoides terciárias em 21% dos tumores de cavidade oral analisados e relataram uma associação entre a presença das estruturas e a menor mortalidade relacionada à doença (WIRSING *et al.*, 2014).

Na caracterização de tumores de cabeça e pescoço realizada pelo TCGA, foram identificados os genes mais frequentemente mutados. Considerando os oito mais frequentes, também foram detectadas frequências similares de mutação na

coorte de 22 amostras do TCGA (Fig 4.14). Por exemplo, os genes *TP53* com 63% de amostras mutadas na coorte de 22 amostras e 72% na coorte inteira e o *FAT1* com 36% de amostras mutadas na coorte de 22 amostras e 23% na coorte inteira. O grupo não metastático possui um maior número de mutações nesses genes, mas também possui uma tendência de maior carga mutacional (Fig 4.15-A). A carga mutacional e de NeoAg possuem uma correlação direta, quanto mais mutação maior o número de NeoAg. Essa tendência é observada na carga de NeoAg, com mediana de NeoAg maior no grupo não metastático (Fig 4.15-B). A cobertura nas regiões intrônicas é menor, o que pode gerar erro de chamada de variantes além dos mecanismos de MMR nas regiões exônicas ser mais ativo que nas intrônicas (FRIGOLA et al., 2017). A predição de NeoAg derivado de TAA leva em consideração a expressão desse TAA e, no grupo metastático, ela possui uma dispersão maior que é refletida na maior dispersão desses NeoAg (Fig 4.15-B, direita).

Outros fatores além da expressão gênica determinam se um peptídeo mutado será um bom NeoAg, como sua afinidade de ligação aos alelos de HLA, eventuais mutações na via de processamento e apresentação de antígeno, a imunogenicidade do p:HLA e a distribuição da mutação que gera o NeoAg na população clonal (SCHUMACHER; SCHEPER; KVISTBORG, 2018). A imunogenicidade do p:HLA, ou seja, o reconhecimento do complexo por um TCR, ainda não pode ser avaliada *in silico* e é um dos principais responsáveis pelos NeoAg não reativos preditos por bioinformática. Os outros fatores conseguem ser considerados pelas ferramentas *in silico*, a filtragem por genes de baixa expressão, por exemplo, foi realizada para excluir peptídeos mutados com poucas chances de apresentação via HLA. A Figura 4.16-A mostra a distribuição da carga de mutação ordenada de forma crescente pela carga de NeoAg. A amostra com maior carga de mutação também é a única na qual foi detectada mutação deletéria em um gene da via de processamento e apresentação, o *PSME2*. Esse gene pertence ao proteassomo, um complexo proteico que degrada proteínas em peptídeos. A menor atividade do proteassomo desfavorece a apresentação de antígenos e pode estar relacionada à maior permissividade de mutações nessa amostra, já que menos NeoAg serão detectados em sua superfície por linfócitos T. Uma das amostras do grupo não metastático destoa das demais, pois possui a terceira maior carga mutacional do grupo, mas

está ordenada como 7^a na carga de NeoAg (Fig 4.16-A; setas). Essa amostra e a anterior (6^a) apresentaram, excepcionalmente, homozigose nos três alelos de HLA-I (dado não mostrado). Isso impacta diretamente nas predições de carga de NeoAg e ajuda a explicar o por quê o número de mutações não acompanha a carga de NeoAg como as outras amostras. Por conta disso, foi verificada a cobertura de leituras na região do HLA-I para descartar possíveis vieses de sequenciamento e foi verificado que, nas 22 amostras, a região estava suficientemente coberta para a predição de alelo pelo Optitype. Uma possível explicação para esse fato é a perda de heterozigose no *loci* do HLA da célula tumoral, o que já foi observado em linhagens de melanoma e em um paciente com leucemia linfóide aguda (PARK et al., 2011; RODRIGUEZ et al., 2005).

A assinatura de transversões e transições (Fig 4.16-C) reflete parte da assinatura mutacional das amostras. O ideal seria identificar o contexto de cada mudança de base, ou seja, qual a base anterior e posterior à base mutada. Isso reflete melhor as alterações de APOBEC, por exemplo. O perfil dominante é o C>T (Citosina para Timina), que está relacionado com a desaminação de citosinas metiladas que resultam em timinas (LUTSENKO; BHAGWAT, 1999). As mutações em genes de reparo ao dano no DNA tendem a aumentar a carga mutacional, o que é observado na Figura 4.16-D. As proteínas ATR, BRCA2 e BRIP1 participam de processos de recombinação homóloga e o PRKDC na via de recombinação não homóloga. Esta última ocorre em casos de quebra da dupla fita de DNA, onde o reparo é realizado na ausência de um DNA molde e está sujeita a erros (SISHC; DAVIS, 2017). Em contrapartida, a recombinação homóloga parte do reparo a partir de um molde (o outro alelo). Mutações em enzimas da via de recombinação homóloga estão associadas à carcinogênese por conta da facilidade em acumular mutações oriundas de erros no reparo e no consequente aumento da instabilidade genômica (BISHOP; SCHIESTL, 2002).

O perfil observado na Figura 4.17-A e B, no qual amostras com maior número de clones de cadeias CDR3 do TCR também possuem maior expressão de moléculas estimulatórias, reforça a possibilidade de maior infiltrado imunológico no tumor. Esse infiltrado parece ser composto também por células supressoras, pois a assinatura de Treg e MDSC bem como a de moléculas inibitórias (e.g. PD1, CTLA-4), é maior nessas amostras. Devido à origem do dado (sequenciamento em

bulk), não há como saber qual população celular é responsável pela expressão de cada molécula. Essas amostras com indícios de maior infiltrado, poderiam se beneficiar da terapia de bloqueio de *checkpoint* imunológico. A regulação positiva de moléculas estimulatórias e inibitórias estiveram acompanhadas na grande maioria dos casos, sugerindo que processos de ativação de linfócitos tenham seu contrabalanço na inibição deste processo por populações e moléculas inibidoras da resposta imunológica. Em contraste com esse grupo, três amostras com regulação negativa dessas moléculas possuíram os menores números de clones de CDR3 (Fig 4.17-C, 1 e 3), o que pode ser um indício de tumores pouco imunogênicos. Dentre as três amostras, uma delas possui mutação em proteína do proteassoma e a menor expressão de *B2M* do grupo. É possível que essa mutação, e sua consequente diminuição na disponibilidades de peptídeos para carregamento nas moléculas de HLA de classe I, tenha algum impacto na regulação da expressão da *B2M*. Contudo, não foram encontradas evidências na literatura que suportem esta regulação direta.

O fato de não terem sido encontradas diferenças estatísticas entre a expressão de *TNF-A*, *IL-6*, *IL-8*, *IFNG*, *IL-10*, *TGFB*, *IDO1*, *TDO*, *TAP1*, *TAP2*, *B2M*, *PDCD1*, *PD-L1*, *PD-L2*, *LAG3*, *TIM-3*, *CTLA4*, *BTLA*, *CD28*, *CD27*, *4-1BB*, *OX40*, *CD40L* e *ICOS* entre os grupos sugere que o desfecho de metástase não depende unicamente da regulação desses genes. Possivelmente a expressão desses genes ocorre em consequência ao microambiente tumoral, seja o tumor metastático ou não.

6 Conclusão

A partir do presente trabalho foi possível caracterizar tumores de cavidade oral com presença ou não de metástase linfonodal. Apesar de diferenças entre as coortes introduzir algum viés nas análises, características de cada grupo podem ser destacadas. Dentre os DEG, a *ADH1B* está mais expressa no grupo metastático e pode indicar uma associação entre o consumo de álcool e esse desfecho. O enriquecimento de duas vias de sinalização envolvendo processos de queratinização no grupo não metastático também pode estar relacionado ao consumo de álcool. Dessa vez, o menor consumo no grupo não metastático explicaria a integridade tecidual da superfície da epiderme, fato a ser explorado em trabalhos futuros.

A análise de populações do microambiente tumoral indicou o enriquecimento de assinaturas de células endoteliais e células-tronco hematopoiéticas em amostras do grupo metastático de ambas as coortes. Por outro lado, células epiteliais, queratinócitos e sebócitos estavam relacionados ao desfecho não metastático, principalmente na coorte ACC. A assinatura de linfócito T CD4 de memória teve forte correlação com o desfecho não metastático em ambas as coortes. Essa célula possui atividade antitumoral, pois fornece estímulos de ativação para CTL e linfócitos B. Tal atividade pode não ter impedido o crescimento tumoral, já que nos dois grupos há tumor, mas a relação dela com o desfecho sugere uma associação no controle de metástases iniciais.

A análise de repertório de TCR e BCR sugere uma maior infiltração de células T no grupo não metastático e de células B no metastático. O perfil mutacional e de NeoAg dos tumores dos dois grupos não apresentou diferença estatística. Contudo, houve uma tendência de maior carga mutacional e de NeoAg no grupo não metastático, possivelmente devido ao tamanho tumoral das amostras deste grupo. Foi possível estabelecer relações sugestivas de escape imunológico entre algumas amostras do grupo não metastático nas quais foram identificadas homozigose dos alelos de HLA e mutações deletérias em proteínas da via de processamento de antígeno. Essas características podem tornar os antígenos tumorais menos perceptíveis ao SI e favorecer o crescimento tumoral. Por outro lado, essas alterações não estiveram tão presentes no grupo metastático, no qual um escape imunológico poderia estar associado à metástase. Nesse caso, parece haver uma

dissociação entre as características de escape imunológico descritas com o escape imunológico favorecendo o desfecho metastático. As populações do microambiente se mostraram diferentes entre os grupos e, ao contrário da carga de NeoAg, elas poderiam ter mais impacto no desfecho de metástase. Um *cluster* de amostras de cada grupo possui um perfil diferente das demais, com regulação positiva de moléculas estimulatórias e inibitórias além de maior número de clones de cadeias de TCR. Isso sugere que mecanismos diferentes de regulação ocorrem intra-grupo independentemente do desfecho metastático. Análises recentes de bioinformática em dados de tumor de cabeça e pescoço do TCGA mostraram o envolvimento do SI nesses tumores. Portanto, o foco em imunoinformática pode ajudar a desvendar esse microambiente complexo.

No presente trabalho, ainda não está totalmente claro qual papel cada componente - NeoAg, populações celulares, repertório T e B - exerce no microambiente tumoral e no favorecimento de metástases. Contudo, alguns indícios que auxiliam a esclarecer esse microambiente foram identificados. Em breve, pretendemos aprimorar nosso *pipeline* de análise e estender a comparação para outros tumores metastáticos e não metastáticos.

7 Referências Bibliográficas

ALMSTEDT, M. et al. The DNA demethylating agent 5-aza-2'-deoxycytidine induces expression of NY-ESO-1 and other cancer/testis antigens in myeloid leukemia cells.

Leukemia Research, v. 34, n. 7, p. 899–905, Jul. 2010.

AMERICAN CANCER SOCIETY. **Risk Factors for Oral Cavity and Oropharyngeal Cancers**. Disponível em:

<<https://www.cancer.org/cancer/oral-cavity-and-oropharyngeal-cancer/causes-risks-prevention/risk-factors.html>>. Acesso em: 17 dec. 2018a.

AMERICAN CANCER SOCIETY. **What Are Oral Cavity and Oropharyngeal Cancers?**

Disponível em:

<<https://www.cancer.org/cancer/oral-cavity-and-oropharyngeal-cancer/about/what-is-oral-cavity-cancer.html>>. Acesso em: 17 dec. 2018b.

AMERICAN CANCER SOCIETY. **Surgery for Oral Cavity and Oropharyngeal Cancer**.

Disponível em:

<<https://www.cancer.org/cancer/oral-cavity-and-oropharyngeal-cancer/treating/surgery.html>>. Acesso em: 25 feb. 2019c.

AOTO, K. et al. Immunogenic tumor cell death induced by chemotherapy in patients with breast cancer and esophageal squamous cell carcinoma. **Oncology Reports**, v. 39, n. 1, p. 151–159, Jan. 2018.

APTE, R. N. et al. Cytokine-induced tumor immunogenicity: endogenous interleukin-1 alpha expressed by fibrosarcoma cells confers reduced tumorigenicity. **Immunology Letters**, v. 39, n. 1, p. 45–52, Dec. 1993.

APTE, R. N. et al. Opposing effects of IL-1 alpha and IL-1 beta on malignancy patterns. Tumor cell-associated IL-1 alpha potentiates anti-tumor immune responses and tumor regression, whereas IL-1 beta potentiates invasiveness. **Advances in Experimental Medicine and Biology**, v. 479, p. 277–288, 2000.

ARAN, D.; HU, Z.; BUTTE, A. J. xCell: digitally portraying the tissue cellular heterogeneity landscape. **Genome Biology**, v. 18, n. 1, p. 220, 15 Nov. 2017.

BAAN, R. et al. Carcinogenicity of alcoholic beverages. **The Lancet Oncology**, v. 8, n. 4, p. 292–293, Apr. 2007.

BABRAHAM BIOINFORMATICS. **FastQC**. Disponível em:

<<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>>. Acesso em: 17 dec. 2018.

BALKWILL, F. Tumor necrosis factor or tumor promoting factor? **Cytokine & growth factor**

- reviews**, v. 13, n. 2, p. 135–141, Apr. 2002.
- BALKWILL, F. TNF-alpha in promotion and progression of cancer. **Cancer metastasis reviews**, v. 25, n. 3, p. 409–416, Sep. 2006.
- BARBER, D. L. et al. Restoring function in exhausted CD8 T cells during chronic viral infection. **Nature**, v. 439, n. 7077, p. 682–687, 9 Feb. 2006.
- BELCHER, R. et al. Current treatment of head and neck squamous cell cancer. **Journal of Surgical Oncology**, v. 110, n. 5, p. 551–574, Oct. 2014.
- BELZ, G. T.; CARBONE, F. R.; HEATH, W. R. Cross-presentation of antigens by dendritic cells. **Critical Reviews in Immunology**, v. 22, n. 5–6, p. 439–448, 2002.
- BENOIT, M.; DESNUES, B.; MEGE, J.-L. Macrophage polarization in bacterial infections. **Journal of Immunology**, v. 181, n. 6, p. 3733–3739, 15 Sep. 2008.
- BISHOP, A. J. R.; SCHIESTL, R. H. Homologous recombination and its role in carcinogenesis. **Journal of Biomedicine & Biotechnology**, v. 2, n. 2, p. 75–85, 2002.
- BLANDER, J. M.; MEDZHITOV, R. Toll-dependent selection of microbial antigens for presentation by dendritic cells. **Nature**, v. 440, n. 7085, p. 808–812, 6 Apr. 2006.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 1 Aug. 2014.
- BOLOTIN, D. A. et al. MiXCR: software for comprehensive adaptive immunity profiling. **Nature Methods**, v. 12, n. 5, p. 380–381, May 2015.
- BOLOTIN, D. A. et al. Antigen receptor repertoire profiling from RNA-seq data. **Nature Biotechnology**, v. 35, n. 10, p. 908–911, 11 Oct. 2017.
- BRADBURY, J. M.; JACKSON, S. P. ATM and ATR. **Current Biology**, v. 13, n. 12, p. R468, 17 Jun. 2003.
- BRAUD, V. M. et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. **Nature**, v. 391, n. 6669, p. 795–799, 19 Feb. 1998.
- BRCKALO, T. et al. Functional analysis of the CD300e receptor in human monocytes and myeloid dendritic cells. **European Journal of Immunology**, v. 40, n. 3, p. 722–732, Mar. 2010.
- BROAD INSTITUTE. **Cancer Genome Analysis**. Disponível em: <https://software.broadinstitute.org/cancer/cga/>. Acesso em: 24 feb. 2019.
- BURNET, F. M. The concept of immunological surveillance. **Progress in experimental tumor research. Fortschritte der experimentellen Tumorforschung. Progres de la**

recherche experimentale des tumeurs, v. 13, p. 1–27, 1970.

BURRELL, R. A. et al. The causes and consequences of genetic heterogeneity in cancer evolution. **Nature**, v. 501, n. 7467, p. 338–345, 19 Sep. 2013.

CANCER GENOME ATLAS NETWORK. Comprehensive molecular characterization of human colon and rectal cancer. **Nature**, v. 487, n. 7407, p. 330–337, 18 Jul. 2012.

CANCER GENOME ATLAS NETWORK. Comprehensive genomic characterization of head and neck squamous cell carcinomas. **Nature**, v. 517, n. 7536, p. 576–582, 29 Jan. 2015.

CANCER RESEARCH UK. **EBV and cancer | Cancer Research UK**. Disponível em: <<https://www.cancerresearchuk.org/about-cancer/causes-of-cancer/infections-hpv-and-cancer/ebv-and-cancer>>. Acesso em: 20 dec. 2018.

CHAE, Y. K. et al. Mutations in DNA repair genes are associated with increased neo-antigen load and activated T cell infiltration in lung adenocarcinoma. **Oncotarget**, v. 9, n. 8, p. 7949–7960, 30 Jan. 2018.

CHAROENTONG, P. et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. **Cell reports**, v. 18, n. 1, p. 248–262, 3 Jan. 2017.

CHAUDHARY, B.; ELKORD, E. Regulatory T cells in the tumor microenvironment and cancer progression: role and therapeutic targeting. **Vaccines**, v. 4, n. 3, 6 Aug. 2016.

CHEN, C.-F. et al. ATR mutations promote the growth of melanoma tumors by modulating the immune microenvironment. **Cell reports**, v. 18, n. 10, p. 2331–2342, 7 Mar. 2017.

CHEN, M.-F. et al. Significance of the TGF- β 1/IL-6 axis in oral cancer. **Clinical Science**, v. 122, n. 10, p. 459–472, 1 May 2012.

CHUANG, C.-Y. et al. Differential impact of IL-10 expression on survival and relapse between HPV16-positive and -negative oral squamous cell carcinomas. **Plos One**, v. 7, n. 10, p. e47541, 31 Oct. 2012.

CLARK, R. E. et al. Direct evidence that leukemic cells present HLA-associated immunogenic peptides derived from the BCR-ABL b3a2 fusion protein. **Blood**, v. 98, n. 10, p. 2887–2893, 15 Nov. 2001.

CORREA, I.; PLUNKETT, T. Update on HER-2 as a target for cancer therapy: HER2/neu peptides as tumour vaccines for T cell recognition. **Breast Cancer Research**, v. 3, n. 6, p. 399–403, 20 Sep. 2001.

COUSSENS, L. M.; WERB, Z. Inflammation and cancer. **Nature**, v. 420, n. 6917, p. 860–867, 26 Dec. 2002.

CRESPO, J. et al. T cell anergy, exhaustion, senescence, and stemness in the tumor microenvironment. **Current Opinion in Immunology**, v. 25, n. 2, p. 214–221, Apr. 2013.

DADMARZ, R. D. et al. Tumor-infiltrating lymphocytes from human ovarian cancer patients recognize autologous tumor in an MHC class II-restricted fashion. **The cancer journal from Scientific American**, v. 2, n. 5, p. 263–272, Oct. 1996.

DE LAS HERAS-RUBIO, A. et al. Ribosomal proteins as novel players in tumorigenesis. **Cancer metastasis reviews**, v. 33, n. 1, p. 115–141, Mar. 2014.

DE SMET, C. et al. The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 93, n. 14, p. 7149–7153, 9 Jul. 1996.

DE SMET, C. et al. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. **Molecular and Cellular Biology**, v. 19, n. 11, p. 7327–7335, Nov. 1999.

DEFFRENNES, V. et al. Constitutive expression of MHC class II genes in melanoma cell lines results from the transcription of class II transactivator abnormally initiated from its B cell-specific promoter. **Journal of Immunology**, v. 167, n. 1, p. 98–106, 1 Jul. 2001.

DEL CAMPO, A. B. et al. Immune escape of cancer cells with beta2-microglobulin loss over the course of metastatic melanoma. **International Journal of Cancer**, v. 134, n. 1, p. 102–113, 1 Jan. 2014.

DOBIN, A. **STAR manual 2.4.0.1**. Disponível em: <<http://labshare.cshl.edu/shares/gingeraslab/www-data/dobin/STAR/STAR.posix/doc/STARmanual.pdf>>. Acesso em: 24 jan. 2019.

DONG, H. et al. Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. **Nature Medicine**, v. 8, n. 8, p. 793–800, Aug. 2002.

DUNN, G. P. et al. Cancer immunoediting: from immunosurveillance to tumor escape. **Nature Immunology**, v. 3, n. 11, p. 991–998, Nov. 2002.

ECSEQ BIOINFORMATICS. **Why does the per base sequence quality decrease over the read in Illumina?** Disponível em: <<https://www.ecseq.com/support/ngs/why-does-the-sequence-quality-decrease-over-the-read-in-illumina>>. Acesso em: 29 jan. 2019.

EHRlich, P. Ueber den jetzigen stand der Karzinomforschung. **Nederlands Tijdschrift Voor Geneeskunde**, v. 5, p. 273–290, 1909.

ELMORE, S. A. Histopathology of the lymph nodes. **Toxicologic Pathology**, v. 34, n. 5, p.

425–454, 2006.

EWING, B. et al. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Research**, v. 8, n. 3, p. 175–185, Mar. 1998.

FARIDI, R. et al. Oncogenic potential of Human Papillomavirus (HPV) and its relation with cervical cancer. **Virology Journal**, v. 8, p. 269, 3 Jun. 2011.

FREMD, C. et al. B cell-regulated immune responses in tumor models and cancer patients. **Oncoimmunology**, v. 2, n. 7, p. e25443, 1 Jul. 2013.

FRIGOLA, J. et al. Reduced mutation rate in exons due to differential mismatch repair. **Nature Genetics**, v. 49, n. 12, p. 1684–1692, Dec. 2017.

GANEM-QUINTANAR, A. et al. Lipid extracting effect on ethanol on keratinized oral mucosa. **Pharmaceutical Research**, v. 15, n. 3, p. 495–498, Mar. 1998.

GAO, J.-X. et al. Targeting PIWIL2-like (PL2L) proteins by a monoclonal antibody for immunotherapy of both solid and hematopoietic cancers. **Journal of Clinical Oncology**, v. 34, n. 15_suppl, p. e14001–e14001, 20 May 2016.

GARDAI, S. J. et al. Cell-surface calreticulin initiates clearance of viable or apoptotic cells through trans-activation of LRP on the phagocyte. **Cell**, v. 123, n. 2, p. 321–334, 21 Oct. 2005.

GASSER, S. et al. The DNA damage pathway regulates innate immune system ligands of the NKG2D receptor. **Nature**, v. 436, n. 7054, p. 1186–1190, 25 Aug. 2005.

GATK. **Phred-scaled Quality Scores — GATK-Forum**. Disponível em: <<https://gatkforums.broadinstitute.org/gatk/discussion/4260/how-should-i-interpret-phred-scaled-quality-scores>>. Acesso em: 3 jan. 2019.

GAUTHIER, J. et al. A brief history of bioinformatics. **Briefings in Bioinformatics**, 3 Aug. 2018.

GERMANO, G. et al. Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth. **Nature**, v. 552, n. 7683, p. 116–120, 7 Dec. 2017.

GETTINGER, S. et al. Impaired HLA class I antigen processing and presentation as a mechanism of acquired resistance to immune checkpoint inhibitors in lung cancer. **Cancer discovery**, v. 7, n. 12, p. 1420–1435, 12 Oct. 2017.

GNJATIC, S. et al. NY-ESO-1: review of an immunogenic tumor antigen. In: *Advances in cancer research*. [s.l.] Elsevier, 2006. v. 95p. 1–30.

GOMES, F. G. et al. Tumor angiogenesis and lymphangiogenesis: tumor/endothelial crosstalk and cellular/microenvironmental signaling mechanisms. **Life Sciences**, v. 92, n. 2,

p. 101–107, 7 Feb. 2013.

GONG, D. et al. TGF β signaling plays a critical role in promoting alternative macrophage activation. **BMC Immunology**, v. 13, p. 31, 15 Jun. 2012.

GUERMONPREZ, P. et al. Antigen presentation and T cell stimulation by dendritic cells. **Annual Review of Immunology**, v. 20, p. 621–667, 2002.

HACKL, H. et al. Computational genomics tools for dissecting tumour-immune cell interactions. **Nature Reviews. Genetics**, v. 17, n. 8, p. 441–458, 4 Jul. 2016.

HADRUP, S.; DONIA, M.; THOR STRATEN, P. Effector CD4 and CD8 T cells and their role in the tumor microenvironment. **Cancer Microenvironment**, v. 6, n. 2, p. 123–133, Aug. 2013.

HANAHAN, D.; WEINBERG, R. A. The hallmarks of cancer. **Cell**, v. 100, n. 1, p. 57–70, 7 Jan. 2000.

HANAHAN, D.; WEINBERG, R. A. Hallmarks of cancer: the next generation. **Cell**, v. 144, n. 5, p. 646–674, 4 Mar. 2011.

HAVELL, E. A.; FIERS, W.; NORTH, R. J. The antitumor function of tumor necrosis factor (TNF), I. Therapeutic action of TNF against an established murine sarcoma is indirect, immunologically dependent, and limited by severe toxicity. **The Journal of Experimental Medicine**, v. 167, n. 3, p. 1067–1085, 1 Mar. 1988.

HEEMSKERK, B.; KVISTBORG, P.; SCHUMACHER, T. N. M. The cancer antigenome. **The EMBO Journal**, v. 32, n. 2, p. 194–203, 23 Jan. 2013.

HEERBOTH, S. et al. EMT and tumor metastasis. **Clinical and translational medicine**, v. 4, p. 6, 26 Feb. 2015.

HERNANDEZ, C.; HUEBENER, P.; SCHWABE, R. F. Damage-associated molecular patterns in cancer: a double-edged sword. **Oncogene**, v. 35, n. 46, p. 5931–5941, 17 Nov. 2016.

HICKLIN, D. J.; MARINCOLA, F. M.; FERRONE, S. HLA class I antigen downregulation in human cancers: T-cell immunotherapy revives an old story. **Molecular medicine today**, v. 5, n. 4, p. 178–186, Apr. 1999.

HODGE, D. R.; HURT, E. M.; FARRAR, W. L. The role of IL-6 and STAT3 in inflammation and cancer. **European Journal of Cancer**, v. 41, n. 16, p. 2502–2512, Nov. 2005.

HOSOI, A. et al. Memory Th1 cells augment tumor-specific CTL following transcutaneous peptide immunization. **Cancer Research**, v. 68, n. 10, p. 3941–3949, 15 May 2008.

HU, P. et al. The prognostic value of cytotoxic T-lymphocyte antigen 4 in cancers: a systematic review and meta-analysis. **Scientific reports**, v. 7, p. 42913, 17 Feb. 2017.

HUMAN PROTEIN ATLAS. **Tissue expression of PIWIL2 - Summary - The Human Protein Atlas**. Disponível em:

<<https://www.proteinatlas.org/ENSG00000197181-PIWIL2/tissue>>. Acesso em: 25 jan. 2019.

IANNELLO, A. et al. p53-dependent chemokine production by senescent tumor cells supports NKG2D-dependent tumor elimination by natural killer cells. **The Journal of Experimental Medicine**, v. 210, n. 10, p. 2057–2069, 23 Sep. 2013.

IARC. **Cancer Today**. Disponível em:

<https://gco.iarc.fr/today/online-analysis-pie?v=2018&mode=cancer&mode_population=continents&population=900&populations=900&key=total&sex=0&cancer=39&type=1&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&nb_items=7&group_cancer=1&include_nmssc=1&include_nmssc_other=1&half_pie=0&donut=0&population_group_globocan_id=>>. Acesso em: 17 dec. 2018.

INCA. **Estimativa 2018**. Disponível em:

<<http://www1.inca.gov.br/estimativa/2018/estimativa-2018.pdf>>. Acesso em: 15 dec. 2018a.

INCA. **O que é câncer?** Disponível em: <<https://www.inca.gov.br/o-que-e-cancer>>. Acesso em: 16 dec. 2018b.

INCA. **Câncer de boca**. Disponível em:

<<https://www.inca.gov.br/tipos-de-cancer/cancer-de-boca>>. Acesso em: 17 dec. 2018c.

JANG, S. J. et al. Activation of melanoma antigen tumor antigens occurs early in lung carcinogenesis. **Cancer Research**, v. 61, n. 21, p. 7959–7963, 1 Nov. 2001.

JOHNSEN, A. K. et al. Deficiency of transporter for antigen presentation (TAP) in tumor cells allows evasion of immune surveillance and increases tumorigenesis. **Journal of Immunology**, v. 163, n. 8, p. 4224–4231, 15 Oct. 1999.

JOHNSON, D. B. et al. Melanoma-specific MHC-II expression represents a tumour-autonomous phenotype and predicts response to anti-PD-1/PD-L1 therapy. **Nature Communications**, v. 7, p. 10582, 29 Jan. 2016.

JOYCE, J. A.; POLLARD, J. W. Microenvironmental regulation of metastasis. **Nature Reviews. Cancer**, v. 9, n. 4, p. 239–252, Apr. 2009.

JURTZ, V. et al. NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. **Journal of Immunology**, v. 199, n. 9, p. 3360–3368, 1 Nov. 2017.

KANG, Y. et al. Dynamic transcriptional regulatory complexes including BORIS, CTCF and Sp1 modulate NY-ESO-1 expression in lung cancer cells. **Oncogene**, v. 26, n. 30, p.

4394–4403, 28 Jun. 2007.

KASAIAN, K.; LI, Y. Y.; JONES, S. J. M. Bioinformatics for cancer genomics. In: **Cancer Genomics**. [s.l.] Elsevier, 2014. p. 133–152.

KELLER, R. et al. Mechanisms of macrophage-mediated tumor cell killing: a comparative analysis of the roles of reactive nitrogen intermediates and tumor necrosis factor. **International Journal of Cancer**, v. 46, n. 4, p. 682–686, 15 Oct. 1990.

KENNEDY, R.; CELIS, E. Multiple roles for CD4+ T cells in anti-tumor immune responses. **Immunological Reviews**, v. 222, p. 129–144, Apr. 2008.

KHONG, H. T.; WANG, Q. J.; ROSENBERG, S. A. Identification of multiple antigens recognized by tumor-infiltrating lymphocytes from a single patient: tumor escape by antigen loss and loss of MHC expression. **Journal of Immunotherapy**, v. 27, n. 3, p. 184–190, Jun. 2004.

KIM, S.-Y. et al. Role of the IL-6-JAK1-STAT3-Oct-4 pathway in the conversion of non-stem cancer cells into cancer stem-like cells. **Cellular Signalling**, v. 25, n. 4, p. 961–969, Apr. 2013.

KOBAYASHI, K. S.; VAN DEN ELSEN, P. J. NLRC5: a key regulator of MHC class I-dependent immune responses. **Nature Reviews. Immunology**, v. 12, n. 12, p. 813–820, Dec. 2012.

KOIZUMI, M. et al. Increased B cell-activating factor promotes tumor invasion and metastasis in human pancreatic cancer. **Plos One**, v. 8, n. 8, p. e71367, 6 Aug. 2013.

KORC, M.; FRIESEL, R. E. The role of fibroblast growth factors in tumor growth. **Current Cancer Drug Targets**, v. 9, n. 5, p. 639–651, 1 Aug. 2009.

KUMAR, V. et al. The Nature of Myeloid-Derived Suppressor Cells in the Tumor Microenvironment. **Trends in Immunology**, v. 37, n. 3, p. 208–220, Mar. 2016.

LANDSKRON, G. et al. Chronic inflammation and cytokines in the tumor microenvironment. **Journal of immunology research**, v. 2014, p. 149185, 13 May 2014.

LANGER, H. F. et al. Complement-mediated inhibition of neovascularization reveals a point of convergence between innate immunity and angiogenesis. **Blood**, v. 116, n. 22, p. 4395–4403, 25 Nov. 2010.

LE, D. T. et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. **Science**, v. 357, n. 6349, p. 409–413, 28 Jul. 2017.

LEE, S. H. et al. TNF α enhances cancer stem cell-like phenotype via Notch-Hes1 activation in oral squamous cell carcinoma cells. **Biochemical and Biophysical Research**

Communications, v. 424, n. 1, p. 58–64, 20 Jul. 2012.

LETTAU, M. et al. Secretory lysosomes and their cargo in T and NK cells. **Immunology Letters**, v. 108, n. 1, p. 10–19, 15 Jan. 2007.

LI, B.; DEWEY, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC Bioinformatics**, v. 12, p. 323, 4 Aug. 2011.

LIEBERMAN, J. The ABCs of granule-mediated cytotoxicity: new weapons in the arsenal. **Nature Reviews. Immunology**, v. 3, n. 5, p. 361–370, May 2003.

LIEPE, J. et al. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. **Science**, v. 354, n. 6310, p. 354–358, 21 Oct. 2016.

LITWIN, M. et al. Aberrant expression of PIWIL1 and PIWIL2 and their clinical significance in ductal breast carcinoma. **Anticancer Research**, v. 38, n. 4, p. 2021–2030, 2018.

LIU, X. V. et al. Ras activation induces expression of Raet1 family NK receptor ligands. **Journal of Immunology**, v. 189, n. 4, p. 1826–1834, 15 Aug. 2012.

LJUNGGREN, H. G.; KÄRRE, K. In search of the “missing self”: MHC molecules and NK cell recognition. **Immunology today**, v. 11, n. 7, p. 237–244, Jul. 1990.

LOVE, M.; ANDERS, S.; HUBER, W. **Analyzing RNA-seq data with DESeq2**. Disponível em: <<http://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/DESeq2.html>>. Acesso em: 17 jan. 2019.

LOVE, M. I.; HUBER, W.; ANDERS, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome Biology**, v. 15, n. 12, p. 550, 2014.

LUTSENKO, E.; BHAGWAT, A. S. Principal causes of hot spots for cytosine to thymine mutations at sites of cytosine methylation in growing cells. A model, its experimental support and implications. **Mutation Research**, v. 437, n. 1, p. 11–20, Jul. 1999.

MAEURER, M. J. et al. Tumor escape from immune recognition: lethal recurrent melanoma in a patient associated with downregulation of the peptide transporter protein TAP-1 and loss of expression of the immunodominant MART-1/Melan-A antigen. **The Journal of Clinical Investigation**, v. 98, n. 7, p. 1633–1641, 1 Oct. 1996.

MARTÍNEZ-LOSTAO, L.; ANEL, A.; PARDO, J. How do cytotoxic lymphocytes kill cancer cells? **Clinical Cancer Research**, v. 21, n. 22, p. 5047–5056, 15 Nov. 2015.

MELCHER, A. et al. Apoptosis or necrosis for tumor immunotherapy: what's in a name? **Journal of Molecular Medicine**, v. 77, n. 12, p. 824–833, Dec. 1999.

MENDEZ, R. et al. HLA and melanoma: multiple alterations in HLA class I and II expression in human melanoma cell lines from ESTDAB cell bank. **Cancer Immunology**,

Immunotherapy, v. 58, n. 9, p. 1507–1515, Sep. 2009.

MIYANISHI, M. et al. Identification of Tim4 as a phosphatidylserine receptor. **Nature**, v. 450, n. 7168, p. 435–439, 15 Nov. 2007.

MONJAZEB, A. M. et al. Immunoediting and antigen loss: overcoming the achilles heel of immunotherapy with antigen non-specific therapies. **Frontiers in oncology**, v. 3, p. 197, 26 Jul. 2013.

MORRIS, E. C. et al. A critical role of T cell antigen receptor-transduced MHC class I-restricted helper T cells in tumor protection. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 22, p. 7934–7939, 31 May 2005.

MORRISON, C. D.; PARVANI, J. G.; SCHIEMANN, W. P. The relevance of the TGF- β Paradox to EMT-MET programs. **Cancer Letters**, v. 341, n. 1, p. 30–40, 28 Nov. 2013.

MOSSER, D. M.; EDWARDS, J. P. Exploring the full spectrum of macrophage activation. **Nature Reviews. Immunology**, v. 8, n. 12, p. 958–969, Dec. 2008.

MUELLER, M. M.; FUSENIG, N. E. Friends or foes - bipolar effects of the tumour stroma in cancer. **Nature Reviews. Cancer**, v. 4, n. 11, p. 839–849, Nov. 2004.

MULTHOFF, G.; MOLLS, M.; RADONS, J. Chronic inflammation in cancer development. **Frontiers in immunology**, v. 2, p. 98, 2011.

MUNN, D. H.; MELLOR, A. L. Indoleamine 2,3-dioxygenase and tumor-induced tolerance. **The Journal of Clinical Investigation**, v. 117, n. 5, p. 1147–1154, May 2007.

MUNN, L. L. Cancer and inflammation. **Wiley interdisciplinary reviews. Systems biology and medicine**, v. 9, n. 2, 2017.

MURPHY, K. The immune system recognizes infection and induces protective responses. In: **Janeway's Immunobiology**. 8th. ed. [s.l.] Garland Science, 2012a. p. 3.

MURPHY, K. The T-cell receptor gene segments are arranged in a similar pattern to immunoglobulin gene segments and are rearranged by the same enzymes. In: **Janeway's Immunobiology**. 8th. ed. [s.l.] Garland Science, 2012b. p. 171.

MURPHY, K. Appendix I - The Immunologist's Toolbox. In: **Janeway's Immunobiology**. 8th. ed. [s.l.] Garland Science, 2012c. p. 717.

MURPHY, K. The MHC class I and class II molecules deliver peptides to the cell surface from two intracellular compartments. In: **Janeway's Immunobiology**. 8th. ed. [s.l.] Garland Science, 2012d. p. 202–203.

MURPHY, K. Many proteins involved in antigen processing and presentation are encoded by genes within the MHC. In: **Janeway's Immunobiology**. 8th. ed. [s.l.] Garland Science,

2012e. p. 217.

MURPHY, K. Newly synthesized MHC class I molecules are retained in the endoplasmic reticulum until they bind a peptide. In: **Janeway's Immunobiology**. 8th. ed. [s.l.] Garland Science, 2012f. p. 207.

MURPHY, K. Rearrangement of V, D, and J gene segments is guided by flanking DNA sequences. In: **Janeway's Immunobiology**. 8th. ed. [s.l.] Garland Science, 2012g. p. 161.

MURPHY, K. Tumor rejection antigens can be recognized by T cells and form the basis of immunotherapies. In: **Janeway's Immunobiology**. 9th. ed. [s.l.] Garland Science, 2017. p. 721–722.

NAGASAKI, T. et al. Interleukin-6 released by colon cancer-associated fibroblasts is critical for tumour angiogenesis: anti-interleukin-6 receptor antibody suppressed angiogenesis and inhibited tumour-stroma interaction. **British Journal of Cancer**, v. 110, n. 2, p. 469–478, 21 Jan. 2014.

NAKANO, Y. et al. Expression of tumor necrosis factor-alpha and interleukin-6 in oral squamous cell carcinoma. **Japanese journal of cancer research : Gann**, v. 90, n. 8, p. 858–866, Aug. 1999.

NATH, S.; MUKHERJEE, P. MUC1: a multifaceted oncoprotein with a key role in cancer progression. **Trends in Molecular Medicine**, v. 20, n. 6, p. 332–342, Jun. 2014.

NATIONAL CANCER INSTITUTE. **Metastatic Cancer - National Cancer Institute**. Disponível em: <<https://www.cancer.gov/types/metastatic-cancer>>. Acesso em: 17 jan. 2019.

NATIONAL CANCER INSTITUTE. **Human Papillomavirus (HPV) Vaccines - National Cancer Institute**. Disponível em: <<https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/hpv-vaccine-fact-sheet>>. Acesso em: 20 dec. 2018.

NAZAROV, V. I. et al. tcR: an R package for T cell receptor repertoire advanced data analysis. **BMC Bioinformatics**, v. 16, p. 175, 28 May 2015.

NEMAZEE, D. Receptor editing in lymphocyte development and central tolerance. **Nature Reviews. Immunology**, v. 6, n. 10, p. 728–740, Oct. 2006.

NETTO, F. O. G. **ESTUDO EM LARGA ESCALA DE ALTERAÇÕES TRANSCRICIONAIS RELACIONADAS À OCORRÊNCIA DE METÁSTASES NO CÂNCER ORAL**. Doctoral dissertation—[s.l.: s.n.].

NEWMAN, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. **Nature Methods**, v. 12, n. 5, p. 453–457, May 2015.

NIH. **Oral Cavity and Pharynx Cancer - Cancer Stat Facts**. Disponível em: <<https://seer.cancer.gov/statfacts/html/oralcav.html>>. Acesso em: 28 jan. 2019.

NIH. **What Is Cancer? - National Cancer Institute**. Disponível em: <<https://www.cancer.gov/about-cancer/understanding/what-is-cancer>>. Acesso em: 16 dec. 2018.

NIH. **Head and Neck Cancers - National Cancer Institute**. Disponível em: <<https://www.cancer.gov/types/head-and-neck/head-neck-fact-sheet>>. Acesso em: 16 dec. 2018.

OFT, M. IL-10: master switch from tumor-promoting inflammation to antitumor immunity. **Cancer immunology research**, v. 2, n. 3, p. 194–199, Mar. 2014.

PAGÈS, F. et al. Immune infiltration in human tumors: a prognostic factor that should not be ignored. **Oncogene**, v. 29, n. 8, p. 1093–1102, 25 Feb. 2010.

PARK, H. et al. False homozygosity results in HLA genotyping due to loss of chromosome 6 in a patient with acute lymphoblastic leukemia. **The Korean journal of laboratory medicine**, v. 31, n. 4, p. 302–306, 3 Oct. 2011.

PARKER, B. S.; RAUTELA, J.; HERTZOG, P. J. Antitumour actions of interferons: implications for cancer therapy. **Nature Reviews. Cancer**, v. 16, n. 3, p. 131–144, Mar. 2016.

PAULSEN, E.-E. et al. CTLA-4 expression in the non-small cell lung cancer patient tumor microenvironment: diverging prognostic impact in primary tumors and lymph node metastases. **Cancer Immunology, Immunotherapy**, v. 66, n. 11, p. 1449–1461, Nov. 2017.

PEGGS, K. S.; QUEZADA, S. A.; ALLISON, J. P. Cancer immunotherapy: co-stimulatory agonists and co-inhibitory antagonists. **Clinical and Experimental Immunology**, v. 157, n. 1, p. 9–19, Jul. 2009.

PELLETIER, J.; THOMAS, G.; VOLAREVIĆ, S. Ribosome biogenesis in cancer: new players and therapeutic avenues. **Nature Reviews. Cancer**, v. 18, n. 1, p. 51–63, Jan. 2018.

PICCININI, A. M.; MIDWOOD, K. S. DAMPening inflammation by modulating TLR signalling. **Mediators of Inflammation**, v. 2010, 13 Jul. 2010.

PIRES, F. R. et al. Oral squamous cell carcinoma: clinicopathological features from 346 cases from a single oral pathology service during an 8-year period. **Journal of applied oral science : revista FOB**, v. 21, n. 5, p. 460–467, Oct. 2013.

POLIMANTI, R.; GELERNTER, J. ADH1B: From alcoholism, natural selection, and cancer to the human phenome. **American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics**, v. 177, n. 2, p. 113–125, Mar. 2018.

PROSKURYAKOV, S. Y.; GABAI, V. L. Mechanisms of tumor cell necrosis. **Current Pharmaceutical Design**, v. 16, n. 1, p. 56–68, Jan. 2010.

RAJASAGI, M. et al. Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. **Blood**, v. 124, n. 3, p. 453–462, 17 Jul. 2014.

RAMMENSEE, H. G.; FRIEDE, T.; STEVANOVIĆ, S. MHC ligands and peptide motifs: first listing. **Immunogenetics**, v. 41, n. 4, p. 178–228, 1995.

RÄSÄNEN, K.; VAHERI, A. Activation of fibroblasts in cancer stroma. **Experimental Cell Research**, v. 316, n. 17, p. 2713–2722, 15 Oct. 2010.

RAULET, D. H.; VANCE, R. E.; MCMAHON, C. W. Regulation of the natural killer cell receptor repertoire. **Annual Review of Immunology**, v. 19, p. 291–330, 2001.

REINHERZ, E. L. $\alpha\beta$ TCR-mediated recognition: relevance to tumor-antigen discovery and cancer immunotherapy. **Cancer immunology research**, v. 3, n. 4, p. 305–312, Apr. 2015.

RESTIFO, N. P. et al. Identification of human cancers deficient in antigen processing. **The Journal of Experimental Medicine**, v. 177, n. 2, p. 265–272, 1 Feb. 1993.

RIZVI, N. A. et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. **Science**, v. 348, n. 6230, p. 124–128, 3 Apr. 2015.

RODRIGUEZ, T. et al. High frequency of homozygosity of the HLA region in melanoma cell lines reveals a pattern compatible with extensive loss of heterozygosity. **Cancer Immunology, Immunotherapy**, v. 54, n. 2, p. 141–148, Feb. 2005.

ROSALES, C. Neutrophil: A Cell with Many Roles in Inflammation or Several Cell Types? **Frontiers in physiology**, v. 9, p. 113, 20 Feb. 2018.

RUDDALL, A. et al. B lymphocytes promote lymphogenous metastasis of lymphoma and melanoma. **Neoplasia**, v. 13, n. 8, p. 748–757, Aug. 2011.

SAHIBZADA, H. A. et al. Salivary IL-8, IL-6 and TNF- α as Potential Diagnostic Biomarkers for Oral Cancer. **Diagnostics (Basel, Switzerland)**, v. 7, n. 2, 9 Apr. 2017.

SASAKI, T. et al. Identification of eight members of the Argonaute family in the human genome. **Genomics**, v. 82, n. 3, p. 323–330, Sep. 2003.

SCANLAN, M. J. et al. Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. **Immunological Reviews**, v. 188, p. 22–32, Oct. 2002.

SCHMIEDER, R.; EDWARDS, R. Quality control and preprocessing of metagenomic datasets. **Bioinformatics**, v. 27, n. 6, p. 863–864, 15 Mar. 2011.

SCHUMACHER, T. N. et al. Peptide selection by MHC class I molecules. **Nature**, v. 350, n.

6320, p. 703–706, 25 Apr. 1991.

SCHUMACHER, T. N.; SCHEPER, W.; KVISTBORG, P. Cancer Neoantigens. **Annual Review of Immunology**, 14 Dec. 2018.

SCHUMACHER, T. N.; SCHREIBER, R. D. Neoantigens in cancer immunotherapy. **Science**, v. 348, n. 6230, p. 69–74, 3 Apr. 2015.

SHANG, B. et al. Prognostic value of tumor-infiltrating FoxP3+ regulatory T cells in cancers: a systematic review and meta-analysis. **Scientific reports**, v. 5, p. 15179, 14 Oct. 2015.

SHARMA, M. et al. Molecular changes in invasive front of oral cancer. **Journal of oral and maxillofacial pathology: JOMFP**, v. 17, n. 2, p. 240–247, May 2013.

SHEVACH, E. M. Mechanisms of foxp3+ T regulatory cell-mediated suppression. **Immunity**, v. 30, n. 5, p. 636–645, May 2009.

SHIGA, K. et al. Cancer-Associated Fibroblasts: Their Characteristics and Their Roles in Tumor Growth. **Cancers**, v. 7, n. 4, p. 2443–2458, 11 Dec. 2015.

SHUKLA, S. A. et al. Cancer-Germline Antigen Expression Discriminates Clinical Outcome to CTLA-4 Blockade. **Cell**, v. 173, n. 3, p. 624–633.e8, 19 Apr. 2018.

SILVERSTEIN, A. M. **A history of immunology**. [s.l.] Elsevier, 2009.

SISHC, B. J.; DAVIS, A. J. The Role of the Core Non-Homologous End Joining Factors in Carcinogenesis and Cancer. **Cancers**, v. 9, n. 7, 6 Jul. 2017.

SOUTHERN ILLINOIS UNIVERSITY. **SIU SOM Histology GI**. Disponível em: <<http://www.siumed.edu/~dking2/erg/oralcav.htm>>. Acesso em: 29 jan. 2019.

SZOLEK, A. et al. OptiType: precision HLA typing from next-generation sequencing data. **Bioinformatics**, v. 30, n. 23, p. 3310–3316, 1 Dec. 2014.

TAO, L. et al. Cancer associated fibroblasts: An essential role in the tumor microenvironment. **Oncology letters**, v. 14, n. 3, p. 2611–2620, Sep. 2017.

THOMAS, L.; LAWRENCE, H. S. Cellular and humoral aspects of the hypersensitive states. **Hoerber-Harper**, p. 529–532, 1959.

THOMPSON, J.; ZIMMERMANN, W. The carcinoembryonic antigen gene family: structure, expression and evolution. **Tumour Biology**, v. 9, n. 2–3, p. 63–83, 1988.

TURAJLIC, S. et al. Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. **The Lancet Oncology**, v. 18, n. 8, p. 1009–1021, 7 Jul. 2017.

UPPSALA UNIVERSITET. Researchers show that genetic background regulates tumor

differences. **ScienceDaily**, 30 Sep. 2015.

VIGNERON, N. Human tumor antigens and cancer immunotherapy. **BioMed research international**, v. 2015, p. 948501, 16 Jun. 2015.

VIGNERON, N. et al. Peptide splicing by the proteasome. **The Journal of Biological Chemistry**, v. 292, n. 51, p. 21170–21179, 22 Dec. 2017.

VORONOV, E. et al. IL-1 is required for tumor invasiveness and angiogenesis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 5, p. 2645–2650, 4 Mar. 2003.

WANG, M. et al. Role of tumor microenvironment in tumorigenesis. **Journal of Cancer**, v. 8, n. 5, p. 761–773, 25 Feb. 2017.

WAUGH, D. J. J.; WILSON, C. The interleukin-8 pathway in cancer. **Clinical Cancer Research**, v. 14, n. 21, p. 6735–6741, 1 Nov. 2008.

WEAGEL, E. et al. Macrophage polarization and its role in cancer. **Journal of clinical & cellular immunology**, v. 06, n. 04, 7 Jul. 2015.

WEBER, J. S. Tumor evasion may occur via expression of regulatory molecules: a case for CTLA-4 in melanoma. **The Journal of Investigative Dermatology**, v. 128, n. 12, p. 2750–2752, Dec. 2008.

WEIR, H. K. et al. Heart Disease and Cancer Deaths - Trends and Projections in the United States, 1969-2020. **Preventing Chronic Disease**, v. 13, p. E157, 17 Nov. 2016.

WEISKOPF, K.; WEISSMAN, I. L. Macrophages are critical effectors of antibody therapies for cancer. **MAbs**, v. 7, n. 2, p. 303–310, 2015.

WHO. **ICD-10 Version:2016**. Disponível em: <<https://icd.who.int/browse10/2016/en#/I10>>. Acesso em: 25 jan. 2019.

WHO. **Cancer**. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/cancer>>. Acesso em: 16 dec. 2018.

WIRSING, A. M. et al. Characterisation and prognostic value of tertiary lymphoid structures in oral squamous cell carcinoma. **BMC clinical pathology**, v. 14, p. 38, 23 Aug. 2014.

WIRTH, T. C.; KÜHNEL, F. Neoantigen Targeting-Dawn of a New Era in Cancer Immunotherapy? **Frontiers in immunology**, v. 8, p. 1848, 19 Dec. 2017.

XI, J. et al. Stimulatory role of interleukin 10 in CD8+ T cells through STATs in gastric cancer. **Tumour Biology**, v. 39, n. 5, p. 1010428317706209, May 2017.

XIANG, X. et al. Prognostic value of PD -L1 expression in patients with primary solid tumors.

Oncotarget, v. 9, n. 4, p. 5058–5072, 12 Jan. 2018.

YAMAGUCHI, H.; CONDEELIS, J. Regulation of the actin cytoskeleton in cancer cell migration and invasion. **Biochimica et Biophysica Acta**, v. 1773, n. 5, p. 642–652, May 2007.

YANG, P. et al. Cancer/testis antigens trigger epithelial-mesenchymal transition and genesis of cancer stem-like cells. **Current Pharmaceutical Design**, v. 21, n. 10, p. 1292–1300, 2015.

YE, Y. et al. Identification of Piwil2-like (PL2L) proteins that promote tumorigenesis. **Plos One**, v. 5, n. 10, p. e13406, 20 Oct. 2010.

YU, A. et al. Presence of lymphocytic infiltrate cytotoxic T lymphocyte CD3+, CD8+, and immunoscore as prognostic marker in patients after radical cystectomy. **Plos One**, v. 13, n. 10, p. e0205746, 11 Oct. 2018.

ZAJAC, P. et al. MAGE-A Antigens and Cancer Immunotherapy. **Frontiers in medicine**, v. 4, p. 18, 8 Mar. 2017.

ZHANG, M. et al. RNA editing derived epitopes function as cancer antigens to elicit immune responses. **Nature Communications**, v. 9, n. 1, p. 3919, 25 Sep. 2018.

ZHAO, S. et al. Testicular defense systems: immune privilege and innate immunity. **Cellular & Molecular Immunology**, v. 11, n. 5, p. 428–437, Sep. 2014.

ZIDAR, N. et al. The role of epithelial-mesenchymal transition in squamous cell carcinoma of the oral cavity. **Virchows Archiv: An International Journal of Pathology**, v. 472, n. 2, p. 237–245, Feb. 2018.

ZITVOGEL, L. et al. Immunogenic tumor cell death for optimal anticancer therapy: the calreticulin exposure pathway. **Clinical Cancer Research**, v. 16, n. 12, p. 3100–3104, 15 Jun. 2010.

ZOU, W. et al. Stromal-derived factor-1 in human tumors recruits and alters the function of plasmacytoid precursor dendritic cells. **Nature Medicine**, v. 7, n. 12, p. 1339–1346, Dec. 2001.

8 Anexos

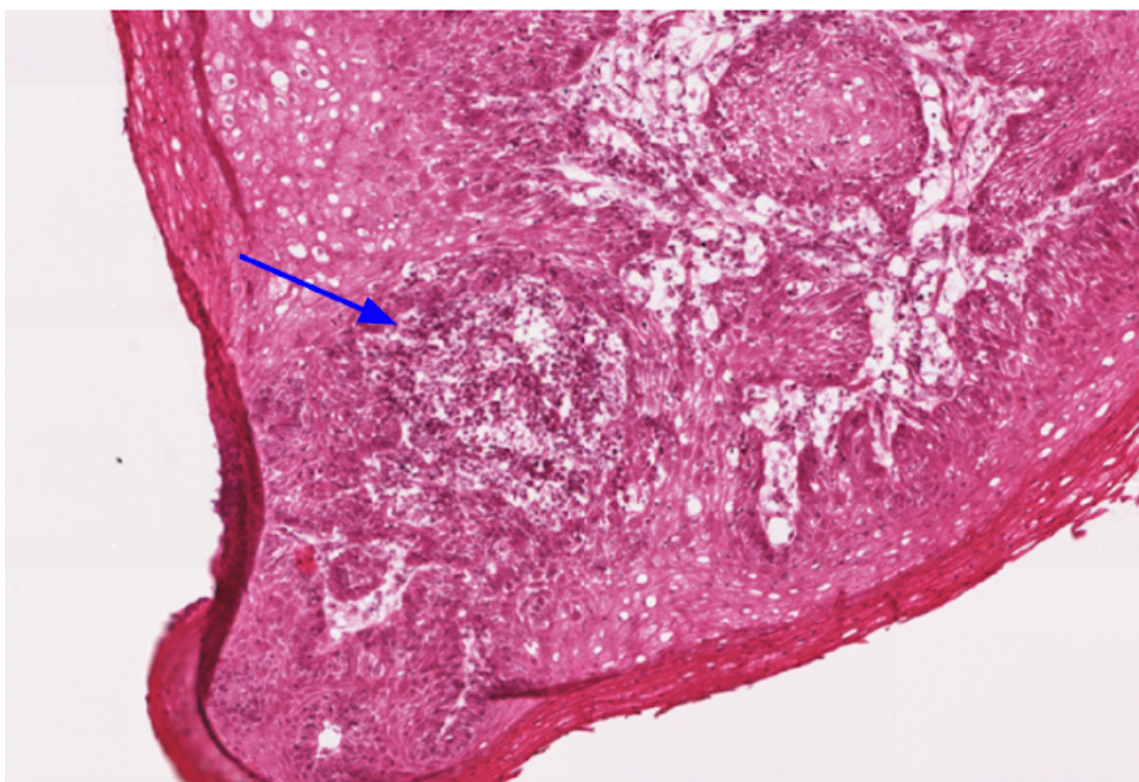
Anexo 1 - Lista dos genes de reparo ao dano no DNA pesquisados

Letra inicial	Gene
A	<i>ABRAXAS1, APEX1, APEX2, ATM, ATR, ATRIP</i>
B	<i>BABAM1, BABAM2, BARD1, BIVM-ERCC5, BLM, BRCA1, BRCA2, BRCC3, BRIP1</i>
C	<i>CCNH, CDK7, CENPS, CENPS-CORT, CENPX, CETN2, CUL4A, CUL4B</i>
D	<i>DCLRE1C, DDB1, DDB2, DNTT</i>
E	<i>EME1, EME2, ERCC1, ERCC2, ERCC3, ERCC4, ERCC5, ERCC6, ERCC8, EXO1</i>
F	<i>FAAP100, FAAP24, FAN1, FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM, FEN1</i>
G	<i>GTF2H1, GTF2H2, GTF2H2C, GTF2H2C_2, GTF2H3, GTF2H4, GTF2H5</i>
H, L	<i>HES1, HMGB1, LIG1, LIG3, LIG4</i>
M	<i>MBD4, MLH1, MLH3, MNAT1, MPG, MRE11, MSH2, MSH3, MSH6, MUS81, MUTYH</i>
N, O	<i>NBN, NEIL1, NEIL2, NEIL3, NHEJ1, NTHL1, OGG1</i>
P	<i>PALB2, PARP1, PARP2, PARP3, PARP4, PAXIP, PCNA, PMS1, PMS2, POLB, POLD1, POLD2, POLD3, POLD4, POLE, POLE2, POLE3, POLE4, POLH, POLI, POLK, POLL, POLM, POLN, PRKDC</i>
R	<i>RAD23A, RAD23B, RAD50, RAD51, RAD51B, RAD51C, RAD51D, RAD52, RAD54B, RAD54L, RBBP8, RBX1, REV1, REV3L, RFC, RFC1, RFC2, RFC3, RFC4, RFC5, RMI1, RMI2, RPA, RPA1, RPA2, RPA3, RPA4</i>
S	<i>SEM1, SLX1A, SLX1B, SLX4, SMUG1, SSBP1, SYCP3</i>
T	<i>TDG, TELO2, TOP3A, TOP3B, TOPBP1</i>
U, W	<i>UBE2T, UIMC1, UNG, USP1, WDR48</i>
X	<i>XPA, XPC, XRCC1, XRCC2, XRCC3, XRCC4, XRCC5, XRCC6</i>

Anexo 2 - Lista dos genes da via de processamento e apresentação pesquisados

Letra inicial	Gene
B	<i>B2M</i>
C	<i>CALR, CANX</i>
E	<i>ERAP1, ERAP2</i>
H	<i>HSPA5</i>
P	<i>PSME1, PSME2, PSME3, PDIA3</i>
T	<i>TAP1, TAP2, TAPBP</i>

Anexo 3 - Lâmina do paciente com estrutura sugestiva de órgão linfoide terciário



Anexo 4 - Carta de aprovação do Comitê de Ética em Pesquisa do AC Camargo Cancer Center



Centro de Tratamento, Ensino e Pesquisa em Câncer

Comitê de Ética em
Pesquisa - CEP

São Paulo, 12 de Dezembro de 2011.

Ao
Dr. Emmanuel Dias-Neto

Ref.: Projeto de Pesquisa nº. 1621/11
"Estudo em larga-escala de alterações transcricionais relacionadas à ocorrência de metástases no câncer oral".

Os membros do Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antonio Prudente – Hospital do Câncer - A.C. Camargo/SP, em sua última reunião de 29/11/2011, **aprovaram** a realização do estudo em referência e tomaram conhecimento dos seguintes documentos:

- Folha de Rosto para Pesquisa Envolvendo Seres Humanos;
- Termo de Compromisso do Pesquisador com as Resoluções do Conselho Nacional de Saúde;
- Termo de Dispensa do Consentimento Livre e Esclarecido;
- Declaração sobre os Dados Coletados, Publicação dos Dados e Propriedade das Informações Geradas;
- Declaração sobre o Uso e Destino do Material Biológico, Publicação dos Dados e Propriedade das Informações Geradas;
- Orçamento Financeiro Detalhado;
- Declaração de Infraestrutura e instalações do Laboratório de Genômica Médica - CIPE;
- Declaração de Ciência e Comprometimento do Biobanco do Hospital A.C. Camargo;
- Declaração sobre a Publicação dos Dados de Pesquisa Utilizando Amostras Fornecidas e Processadas pelo Biobanco do Hospital A.C. Camargo;
- Declaração de Ciência e Comprometimento da Faculdade de Medicina de São José do Rio Preto;
- Declaração de Ciência e Comprometimento do Departamento de Cirurgia de Cabeça e Pescoço e Otorrinolaringologia.

Informações a respeito do andamento do referido projeto deverão ser encaminhadas à assistente do CEP dentro de 12 meses.

Atenciosamente,

Dr. Jefferson Luiz Gross

1º Vice-Coordenador do Comitê de Ética em Pesquisa

1/1

Fundação Antonio Prudente - CNPJ/MF N. 60.961.968/0001-06
Rua Prof. Antônio Prudente, 211 - Liberdade - São Paulo, SP - 01509-900
Telefone: (11) 2189-5000
www.accamargo.org.br