

**INVESTIGAÇÃO DE ASSINATURAS  
MUTACIONAIS EM  
ADENOCARCINOMAS GÁSTRICOS  
E AVALIAÇÃO DE SEU POSSÍVEL  
IMPACTO PROGNÓSTICO**

**MONIZE NAKAMOTO PROVVISOR  
SANTOS**

Dissertação de mestrado apresentada a  
Fundação Antônio Prudente para a obten-  
ção do título de Mestre em Ciências.

Área de concentração: Oncologia

Orientador: Prof. Dr. Israel Tojal da Silva

São Paulo

2018

## FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Santos, Monize Nakamoto Provisor

**Investigação de assinaturas mutacionais em adenocarcinomas gástricos e avaliação de seu possível impacto prognóstico** / Monize Nakamoto Provisor

Santos - São Paulo, 2018

51p.

Dissertação (Mestrado)-Fundação Antônio Prudente.

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

Orientador: Israel Tojal da Silva

Descritores: 1. Biologia Computacional/Computational Biology. 2. Neoplasias Gástricas/Stomach Neoplasms. 3. Prognóstico/Prognosis. 4. Instabilidade de Microssatélites/Microsatellite Instability. 5. Reparo de Erro de Pareamento de DNA/DNA Mismatch Repair

*"We can judge our progress by the courage of our questions and the depth of our answers, our willingness to embrace what is true rather than what feels good."*

*Carl Sagan*

*Este trabalho é dedicado à minha família*

*Por todo carinho, amor e apoio que transmitiram em minha vida..*

## AGRADECIMENTOS

À minha família por toda força concedida para superar os obstáculos que surgiram durante meu trajeto e por estarem sempre presentes.

Ao meu orientador, Dr. Israel Tojal da Silva, pela confiança, orientação, incentivo, conselhos e dedicação nesses anos. A oportunidade de fazer parte de sua equipe, sua ajuda em momentos importantes e seus ensinamentos possibilitaram a realização de uma conquista pessoal de extrema importância em minha vida.

Ao grupo LBCB pela amizade, companheirismo e pelos momentos de alegria. Muito obrigada pela paciência, pelos ensinamentos e por serem fundamentais para a minha formação.

À Fundação Antônio Prudente por todo suporte e ensino.

Ao Dr. Emmanuel Dias-Neto, Dra Diana Noronha Nunes, Dr. Vladmir Cláudio Cordeiro de Lima, Dr. Joao Paulo da Silveira Nogueira Lima, Dr. Helano Carioca Freitas e Dr. Kenneth John Gollob pelos ensinamentos, pela oportunidade de compartilhar e discutir meus achados e por terem sido fundamentais para o desenvolvimento desse estudo.

À Jennice, Mayara, Marianna, Ana Flávia, Luiza, Matheus, Fabio, Danielle e Felipe pela amizade, companheirismo e pelos momentos de alegria compartilhados nesses anos.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP),

por meio do processo nº 2016/11791-7, pela concessão da bolsa de mestrado.

## RESUMO

Provisor M. **Investigação de assinaturas mutacionais em adenocarcinomas gástricos e avaliação de Seu possível impacto prognóstico.** São Paulo; 2018. [Dissertação de mestrado - Fundação Antônio Prudente]

Adenocarcinomas gástricos (**AdGs**) são neoplasias de grande agressividade e alta incidência no Brasil. Os altos índices mundiais de mortalidade dos **AdGs** se devem aos poucos sinais clínicos da doença em seus estágios iniciais, o que leva a uma detecção tardia na maioria dos casos. Sendo o **AdG** uma doença de forte componente genético e ambiental em sua etiologia, a busca por processos causadores de mutações somáticas que contribuem para a sua formação e progressão pode permitir avanços no prognóstico e na determinação de abordagens terapêuticas. Devido à escassez de estudos que associam características moleculares à clínica, além da inexistência de dados moleculares em grande escala obtidos da população brasileira, este trabalho teve por objetivo analisar o perfil mutacional de amostras de **AdG** em busca de assinaturas moleculares que possam ser associadas a aspectos clínicos de interesse. Foram identificadas seis assinaturas mutacionais, através do *signeR*, para uma coorte de 756 amostras. O subtipo molecular **MSI**, a ausência de metástase linfonodal (N0) e de recidiva (*Tumor Free*) demonstraram associação com três assinaturas *signeR*-S1, *signeR*-S2 e *signeR*-S4, associadas aos processos de falha no mecanismo de reparo de DNA (via *MMR*) e desaminação espontânea da 5-metil-citosina. De maneira complementar, este grupo de assinaturas também demonstraram-se associados com uma melhor sobrevida, maior número de células T CD8+, além de uma expressão aumentada de genes de perfil citotóxico. Estes resultados, em conjunto, sugerem que o grupo de pacientes enriquecidos para estas assinaturas, possuem um contexto imunológico favorável à imunoterapia.

## SUMMARY

Provisor M. [**Investigation of mutational signatures in gastric adenocarcinoma and evaluation of its possible prognostic impact**]. São Paulo; 2018. [Dissertação de mestrado - Fundação Antônio Prudente]

Gastric Adenocarcinomas (**GAs**) are highly aggressive neoplasms with a high incidence in Brazil. The high world mortality rates of **GA** are due to few clinical signs of the disease in its early stages which lead to late detection. Because of a strong genetic and environmental component in **GA** etiology, the search for the causative somatic mutations processes that contribute to their initiation and progression may allow advances in prognosis and in the establishment of therapeutic approaches. Due to the lack of studies that associate molecular characteristics with clinical features, as well as the lack of large scale molecular studies in the Brazilian population, this study aims to analyze the mutational profile of **GAs** samples from Brazilian and global populations, looking for molecular signatures that may be associated with relevant clinical aspects. The **signeR** algorithm detected six mutational signatures in a cohort of 756 samples. Three signatures (**signeR-S1 S2 and S4**) were associated with failure in DNA repair mechanisms (MMR pathway) and spontaneous deamination of 5-methylcytosine and had correlation with microsatellite instability (MSI) molecular subtype. Those signatures were also associated with better clinical outcomes like the absence of lymph node metastasis (N0) and longer disease-free survival. Furthermore, those signatures were also associated with a particular lymphocyte infiltration pattern, enriched for T CD8+ lymphocytes, and a higher expression of cytotoxic genes. Together, these results suggest that the group of patients enriched for these signatures may have a favorable prognosis due to better anti tumoral effector immune response and that could be further improved with immunotherapy.



## LISTA DE FIGURAS

Figura 1 – Processos mutacionais operantes no câncer . . . . .	4
Figura 2 – Classificação molecular de Adenocarcinoma Gástrico . . . . .	8
Figura 3 – Mecanismo de bloqueio de <i>checkpoint</i> imune . . . . .	13
Figura 4 – Representação da matriz de mutações . . . . .	24
Figura 5 – Distribuição das mutações nos subtipos moleculares . . . . .	25
Figura 6 – Conjunto de 6 assinaturas mutacionais . . . . .	26
Figura 7 – Similaridade entre as assinaturas estimadas pelos <i>signeR</i> e <i>COSMIC</i> . . . . .	28
Figura 8 – Exposição diferencial das assinaturas e perfil clínico . . . . .	29
Figura 9 – Curvas de sobrevida global dos pacientes nos grupos <b>B</b> e <b>T</b> . . . . .	30
Figura 10 – Distribuição acumulada da exposição de S1S4 nas três coortes . . . . .	32
Figura 11 – Distribuição do <i>status MSI</i> nas amostras do grupo T e B para S1S4 . . . . .	33
Figura 12 – Perfil das células B, TCD4 e TCD8 nos grupos <b>B</b> e <b>T</b> . . . . .	35
Figura 13 – Diferença do nível de expressão de genes marcadores de cito- toxicidade . . . . .	35
Figura 14 – Diferença do nível de expressão de genes marcadores de regulação . . . . .	36
Figura 15 – Expressão dos genes marcadores de citotoxicidade/ativação e CD8+ . . . . .	37
Figura 16 – Correlação da expressão de <i>PD-L1</i> e <i>score</i> de pureza e imune . . . . .	38
Figura 17 – Distribuição da composição do infiltrado nos grupos <b>B</b> e <b>T</b> . . . . .	39
Figura 18 – Conjunto de assinaturas estimadas pelo método <i>signeR</i> . . . . .	40
Figura 19 – <i>Heatmap</i> da similaridade entre as assinaturas . . . . .	41
Figura 20 – Curvas de da sobrevida estimadas através método de <i>Kaplan- Meier</i> . . . . .	42
Figura 21 – Classificação do <i>status MSI</i> nos grupos <b>B</b> e <b>T</b> . . . . .	43

## LISTA DE TABELAS

Tabela 1 – Distribuição das características clínicas dos pacientes no grupo T 31

Tabela 2 – Distribuição das características clínicas dos pacientes do grupo B 31

## LISTA DE SIGLAS E ABREVIATURAS

<b>AdG</b>	Adenocarcinoma Gástrico
<b>CIN</b>	Cromossomicamente Instável
<b>CG</b>	Câncer Gástrico
<b>CNV</b>	Alteração do Número de Cópias
<b>EBV</b>	vírus <i>Epstein-Barr</i>
<b>GS</b>	Genomicamente Estável
<b>MMR</b>	Falha na via de Reparo
<b>MSI</b>	Instabilidade por Microsatélites
<b>MSI-H</b>	ALto nível de Instabilidade por Microsatélites
<b>MSI-L</b>	Baixo nível de Instabilidade por Microsatélites
<b>MSS</b>	Estabilidade por Microsatélites
<b>NGS</b>	Sequenciamento de Nova Geração
<b>NMF</b>	Fatorização Binomial Não-Negativa
<b>PCR</b>	Reação em Cadeia de Polimerase
<b>SNV</b>	Substituições de Única Base

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
1.1	Câncer e Assinaturas Mutacionais	3
1.2	Câncer Gástrico	6
1.3	Adenocarcinoma Gástrico	7
1.3.1	Classificação Histológica	7
1.3.2	Classificação Molecular	8
1.3.3	Instabilidade por Microssatélites (MSI)	9
1.3.4	MSI e prognóstico	10
1.3.5	MSI e Imunoterapia	11
<b>2</b>	<b>OBJETIVOS</b>	<b>16</b>
2.1	Objetivo geral	16
2.2	Objetivos específicos	16
<b>3</b>	<b>METODOLOGIA</b>	<b>17</b>
3.1	Dados moleculares e clínicos	17
3.2	Identificação de Assinaturas Mutacionais	17
3.3	Associação entre as Assinaturas Mutacionais e o prognóstico	18
3.4	Estratificação de pacientes	18
3.5	Associação das Assinaturas Mutacionais e sobrevida	19
3.6	Associação das assinaturas às mutações em genes das vias de reparo de DNA	19
3.7	Exposição da assinatura em diferentes coortes	20
3.8	Assinaturas mutacionais e composição do infiltrado inflamatório	20
3.9	Avaliação do status MSI	21
3.10	Análise da expressão de <i>PD-L1</i> no infiltrado inflamatório e no tumor	22

3.11	Análise da coorte de validação . . . . .	22
<b>4</b>	<b>RESULTADOS E DISCUSSÃO . . . . .</b>	<b>24</b>
4.1	Análise do perfil mutacional das amostras . . . . .	24
4.2	Identificação das Assinaturas Mutacionais . . . . .	25
4.3	Associação das assinaturas ao prognóstico e sobrevida . . . . .	27
4.4	Caracterização clínica dos pacientes dos grupos T e B para assina- turas marcadoras de bom prognóstico . . . . .	30
4.5	Mutações em genes de vias de reparo de DNA . . . . .	31
4.6	Contribuição das coortes públicas para as assinaturas de interesse prognóstico . . . . .	32
4.6.1	Classificação da Instabilidade de Microsatélite ( <i>MSIseq</i> ) . . . . .	32
4.7	Associação das assinaturas com perfil imunológico . . . . .	34
4.8	Análise da expressão de <i>PD-L1</i> no infiltrado e no tumor . . . . .	36
4.9	Análises da coorte de validação . . . . .	38
4.9.1	Assinaturas e Processos Mutacionais . . . . .	38
4.9.2	Análise de sobrevida . . . . .	39
4.9.3	Classificação de Instabilidade por Microsatélite ( <i>MSIseq</i> ) . . . . .	40
<b>5</b>	<b>CONCLUSÕES . . . . .</b>	<b>44</b>
<b>6</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS . . . . .</b>	<b>45</b>

## 1 INTRODUÇÃO

O câncer gástrico (**CG**) é uma doença agressiva com alto impacto global na saúde pública (FERRO et al. 2014), sendo a segunda causa de morte por câncer em nível mundial (JEMAL et al. 2011; FERRO et al. 2014). A falta de métodos consolidados no rastreamento da doença precoce e a baixa responsividade a tratamento curativo são fatores contribuintes para o prognóstico reservado do paciente com câncer de estômago (MACONI et al. 2008).

O Adenocarcinoma Gástrico (**AdG**) é o subtipo histológico mais frequente de **CG** e caracteriza-se por ser heterogêneo e agressivo (JEMAL et al. 2011; FERRO et al. 2014), sendo responsável por acometer cerca de 95% dos casos (ANEIROS-FERNANDEZ et al. 2010). Histologicamente, o **AdG** foi classificado por LAUREN (1965) em dois grupos principais: intestinal e difuso, que possuem perfis clínicos distintos. O subtipo intestinal apresenta células com alto grau de diferenciação, sendo mais comum em populações de alto risco, como é o caso das populações de origem asiática (MARRELLI et al. 2002). Em contrapartida, o subtipo difuso apresenta células com menor grau de diferenciação, com perfil mais invasivo e disseminação local e à distância mais precoce (YAMASHITA et al. 2009).

Estudos moleculares atuais estão se valendo de melhorias das técnicas e redução de custo conseguiram explorar com maior resolução as características moleculares de **AdG** visando traçar perfis de potencial utilidade clínica (BASS et al. 2014). BASS et al. (2014) realizaram uma avaliação extensa dos aspectos moleculares de 295 **AdG** primários, propondo uma classificação molecular em quatro subtipos: i) tumores positivos para o vírus *Epstein-Barr* (do inglês **EBV+**, *Epstein-Barr virus*); ii) tumores com instabilidade de microssatélites (do inglês **MSI**, *Microsatellite Instability*); iii) tumores genomicamente estáveis (do inglês **GS**, *Genomically Stable*); e iv) tumores cromossomicamente instáveis

(**CIN**, *Chromosomal Instability*).

O grupo **MSI** é caracterizado por alterações de inativação genética e epigenética de genes associados ao reparo de DNA, possui forte relação com a carcinogênese (YAMAMOTO et al. 2012). Este perfil molecular, presente em diversas neoplasias hereditárias e esporádicas, vem sendo utilizado como fator prognóstico em vários tipos de câncer (POPAT et al. 2005). Em câncer colorretal e de endométrio, tumores classificados como **MSI** associam-se a um bom prognóstico (GUASTADISEGNI et al. 2010). No entanto, o papel deste perfil molecular no prognóstico de câncer gástrico ainda permanece inconclusivo (OKI et al. 2009; CHOI et al. 2014; ZHU et al. 2015).

Uma vez que o **CG** é uma doença complexa, a identificação dos processos causadores de mutações somáticas que contribuem para a sua formação e progressão são de particular interesse e alvo de investigação (BALMAIN 2002; ALEXANDROV e STRATTON 2014). Fatores endógenos e exógenos, tais como falhas no mecanismo de reparo do DNA, variações no DNA resultantes da ação de enzimas de edição de ácidos nucleicos, tabagismo e exposição à radiação ultravioleta, além do papel de certos componentes da dieta, são reconhecidos como processos causadores de mutações somáticas (STRATTON 2011). Ao agir sobre o genoma, estes processos levam a diferentes combinações de mutações, conhecidas como assinaturas mutacionais, que podem auxiliar na compreensão da biologia do câncer em um contexto evolutivo (HELLEDAY et al. 2014).

Os estudos realizados até o momento que propuseram explorar os aspectos moleculares em larga escala e epidemiológicos foram restritos às populações da Europa, América do Norte e, principalmente, da Ásia (KIM et al. 2010; DUELL et al. 2011; EVERATT et al. 2012). Diante da grande diversidade genética da nossa população e das distintas características clínico-epidemiológicas dos subtipos tumorais, estudos genômicos voltados à detecção de mutações e de seus padrões de ocorrência na população brasileira são fundamentais para ampliar o conhecimento sobre o espectro mutacional envolvido na patologia, bem como

sobre os mecanismos regulatórios envolvidos na progressão tumoral.

## 1.1 CÂNCER E ASSINATURAS MUTACIONAIS

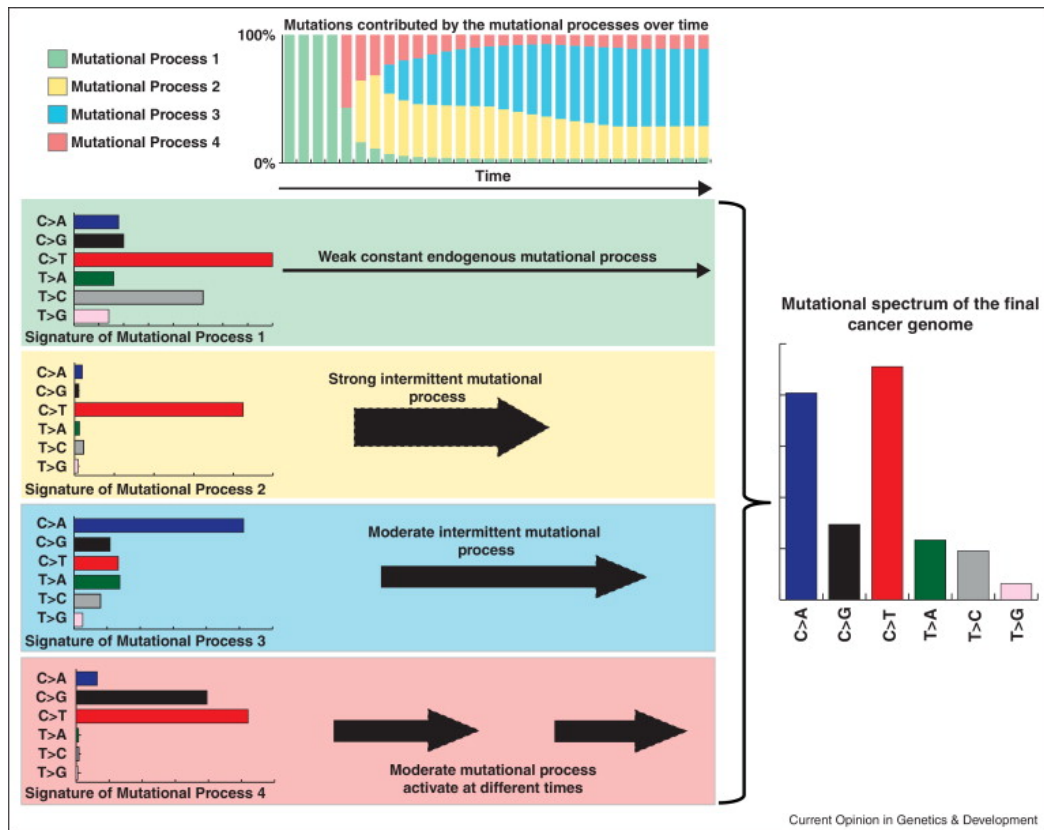
Atualmente, é amplamente aceito que o câncer é o resultado do acúmulo gradual de mutações em genes drivers que levam a um aumento da proliferação celular (HANAHAN e WEINBERG 2011). Mas quais os mecanismos que causam essas mutações? Alguns processos causadores de mutações somáticas não são totalmente conhecidos e ainda são alvos de intensa investigação.

Fatores endógenos e exógenos, tais como falha no mecanismo de reparo do DNA, variações no DNA resultantes da ação de enzimas de edição de ácidos nucléicos, tabagismo, exposição à radiação ultravioleta e hábitos alimentares e exposição ocupacional são reconhecidos como alguns elementos causadores de mutações somáticas (STRATTON 2011; ALEXANDROV et al. 2013).

O espectro de mutações que podem ocorrer no genoma é amplo e compreende inserções, deleções, alteração do número de cópias (CNV), rearranjos e substituições de uma única base (SNV, *Single Nucleotide Variation*) (STRATTON et al. 2009). SNV é o tipo de variação mais comum e pode ser mapeado em seis possíveis trocas: C:G>A:T, C:G>G:C, C:G>T:A, T:A>A:T, T:A> C:G e T:A>G:C. Ao considerar uma base adjacente a cada uma das 6 substituições, um total de 96 trincas podem ser geradas. As combinações únicas e a frequência dessas trincas refletem assinaturas resultante de processos mutacionais que estão ativos no tumor (Figura 1).

NIK-ZAINAL et al. (2012) utilizou a combinação e a frequência dessas trincas, em conjunto com uma generalização do método NMF (*Non-negative Matrix Factorization*) para extrair 5 assinaturas mutacionais (A,B,C,D e E) de 21 genomas de pacientes com câncer de mama. A assinatura A associou-se ao processo de desaminação espontânea de 5-metilcitosina, caracterizando-se por transições no contexto de C>T. A assinatura B, responsável pela maior parte das





Fonte: STRATTON et al. (2009).

**Figura 1** – Processos mutacionais operantes no câncer. Processos mutacionais distintos podem operar ao longo da vida celular. A combinação e a frequência das bases no contexto das seis classes de possíveis substituições somáticas dá origem à uma única assinatura mutacional.

mutações nas amostras deste estudo, encontrava-se presente em cerca de 10% dos tumores de mama ER positivos (STEPHENS et al. 2012) e caracterizava-se por substituições de bases no contexto C>T, C>G e C>A. As demais assinaturas identificadas (C, D e E) ainda permanecem com etiologia desconhecida.

O repertório de assinaturas mutacionais foi expandido por ALEXANDROV et al. (2013), onde um conjunto de 21 assinaturas mutacionais<sup>1</sup> foram detectadas. Neste estudo, foi realizada uma análise ampla de cerca de 5 milhões de *SNVs* provenientes de 7.072 tumores primários disponíveis no TCGA e outros bancos. Essas assinaturas, por sua vez, encontravam-se associadas a processos sem etiologia conhecida e a outros já conhecidos, tais como: idade, falha no mecanismo de reparo de DNA, radiação UV (ultra-violeta), tabagismo e ação das

<sup>1</sup> <https://cancer.sanger.ac.uk/cosmic/signatures>

enzimas da família APOBEC (*Apolipoprotein B mRNA Editing Enzyme Catalytic Polypeptide-like*), dentre outros. Este estudo revelou duas assinaturas que estão associadas à atividade das enzimas APOBEC e, interessante, a desregulação dessas enzimas pode ser uma fonte de instabilidade genômica ao induzir mutações no DNA, já observada em vários tumores. Assim, a detecção de assinaturas permite a compreensão a respeito da etiologia dos processos mutacionais que estão operando no tumor.

Uma vez que o câncer é movido pelo acúmulo de alterações genéticas no DNA genômico de células somáticas, a medicina baseada em informações individuais sobre o genoma tumoral tem sido vista como promessa para o futuro tratamento do câncer (COLLINS e VARMUS 2015). Apesar de alguns estudos relatarem a alta carga mutacional como um fator associado com maior chance de desencadeamento de uma resposta imune antitumoral, outros apontam que tipos específicos de mutações poderiam ser fatores prognósticos mais eficazes na predição de resposta à imunoterápicos (HUGO et al. 2017; SMID et al. 2016).

Uma forma de caracterizar o perfil mutacional de um conjunto de amostras é através da busca por assinaturas mutacionais. Este tipo de abordagem foi validado por WANG et al., que identificaram, de forma inédita, a assinatura mutacional de APOBEC como marcadora e preditiva de resposta à imunoterapia em tumores de pulmão de células não-pequenas. Embora alguns métodos para a detecção de assinatura estejam disponíveis (GEHRING et al. 2015; CARIO e WITTE 2018; LEE et al. 2018), ainda é necessário a supervisão do pesquisador, a fim de definir *a priori* o total de processos agindo nos tumores. A correta estimação do número de assinaturas é um dos desafios e poucas abordagens têm buscado uma solução adequada para esta questão (FISCHER et al. 2013; ROSALES et al. 2017).

## 1.2 CÂNCER GÁSTRICO

O Câncer Gástrico (**CG**) é uma doença agressiva com importante impacto global na saúde pública (FERRO et al. 2014). Agentes infecciosos, incluindo a bactéria *Helicobacter pylori* (*H. pylori*) (UEMURA et al. 2001) e o vírus *Epstein Barr* (**EBV**) (MURPHY et al. 2009), além de fatores ambientais, como dietas ricas em sal e tabagismo, são tidos como importantes fatores de risco para o desenvolvimento desta doença (SHAH e AJANI 2010).

Apesar da diminuição considerável do número de casos nos últimos anos devido à maior disponibilidade de refrigeração para a preservação de alimentos, bem como aos programas de erradicação de *H. pylori*, o **CG** ainda é a segunda maior causa de morte por câncer no mundo (JEMAL et al. 2011; FERRO et al. 2014). No Brasil, são esperados mais de 20.000 novos casos em 2016, com altas taxas de incidência para ambos os sexos (13,0/100.000 em homens e 7,4/100.000 em mulheres) (Ministério da Saúde. Instituto Nacional de Câncer José Alencar Gomes da Silvaa. 2016). Além disso, sem considerar os tumores de pele não-melanoma, o **CG** é o segundo tipo mais incidente em homens das regiões Norte (11,1/100 mil) e Nordeste (10,3/100 mil) (Ministério da Saúde. Instituto Nacional de Câncer José Alencar Gomes da Silva. 2014).

A relação casos novos/mortalidade no **CG** é superior às neoplasias mais incidentes no mundo onde, a cada ano, para cada 800 mil pessoas diagnosticadas com a doença, cerca de 723 mil vão à óbito (FERLAY et al. 2015). O principal fator contribuinte para esta alta taxa de mortalidade consiste na dificuldade de detecção precoce da doença, uma vez que os sintomas clínicos surgem somente em estágios mais avançados, reduzindo as possibilidades de tratamento curativo e, conseqüentemente, as taxas de sobrevida (MACONI et al. 2008).

Histologicamente, o **CG** encontra-se subdividido em três tipos principais, sendo o Adenocarcinoma Gástrico (**AdG**) sua forma mais frequente e responsável por 95% dos casos (ANEIROS-FERNANDEZ et al. 2010).

## 1.3 ADENOCARCINOMA GÁSTRICO

### 1.3.1 Classificação Histológica

O **AdG** caracteriza-se por ser um tipo heterogêneo e agressivo de **CG** (FERRO et al. 2014; JEMAL et al. 2011). A classificação histológica proposta por LAUREN baseia-se em aspectos patológicos e clínicos e é bastante utilizada na prática médica. Nessa classificação, há três grandes grupos : intestinal e difuso, além do subtipo misto, que possui componentes de ambos (HWANG et al. 2010).

O **AdG** de subtipo intestinal é caracterizado por possuir células com alto grau de diferenciação, sendo mais comum em populações de alto risco, como é o caso das populações de origem asiática (MARRELLI et al. 2002). Por outro lado, o tipo difuso é caracterizado por possuir células com menor grau de diferenciação, com perfil mais invasivo e capacidade de se infiltrar em toda a parede gástrica (YAMASHITA et al. 2009).

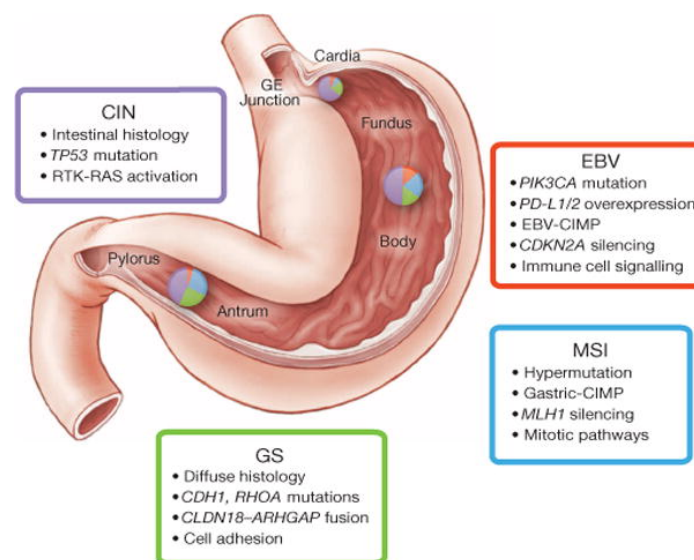
Em 2010, a OMS propôs uma nova classificação baseada nos padrões histológicos predominantes, que coexistem frequentemente com elementos menos dominantes. De acordo com este critério, foram definidos quatro padrões histológicos principais de **CG**: tubular, papilar, mucinoso e pouco coesivo, além de variantes histológicas incomuns (SAKURAI et al. 2014).

O subtipo tubular é o mais comum, sendo relatado mais frequentemente em estágios patológicos iniciais. Histologicamente, caracteriza-se por apresentar túbulos irregulares, fundidos ou ramificados de diferentes tamanhos, com frequente muco intraluminal, restos nucleares e inflamatórios. O subtipo papilar, também mais frequentemente relatado no **CG** precoce, tende a afetar pessoas mais velhas e está associada a altas taxas de comprometimento linfonodal. Histologicamente, caracteriza-se por projeções epiteliais sustentadas por um núcleo fibrovascular central. O subtipo mucinoso, relatado em aproximadamente 10% dos **CG**, caracteriza-se por apresentar células irregulares de arquitetura glan-

dular, com agregados mucinosos extracelulares que constituem pelo menos 50% do volume tumoral. O carcinoma de células em anéis de sinete e os demais carcinomas de células pouco diferenciadas possuem células que se assemelham morfológicamente a histiócitos, linfócitos e plasmócitos. Estes subtipos possuem maior propensão à invasão linfovascular e metástase linfonodal (SAKURAI et al. 2014).

### 1.3.2 Classificação Molecular

Em 2014, BASS et al., em conjunto com pesquisadores do grupo TCGA (*The Cancer Genome Atlas*), apresentaram uma nova classificação molecular com potencial utilidade clínica (Figura 2). Neste estudo, foram sequenciados o exoma (WES) de 295 amostras pareadas (normal e tumor) de CG. A análise deste número considerável de amostras só foi possível graças à maior acessibilidade às técnicas de sequenciamento de última geração (NGS, *Next Generation Sequencing*) e da expressiva redução nos custos destes protocolos.



Fonte: BASS et al. (2014).

**Figura 2** – Classificação molecular de Adenocarcinoma Gástrico. Esta imagem lista as principais características associadas a cada um dos quatro subtipos moleculares estabelecidos para câncer gástrico.

BASS et al. (2014) propuseram quatro subtipos moleculares: i) tumores

positivos para o vírus *Epstein-Barr* (**EBV+**), caracterizados por apresentarem mutações frequentes no gene *PIK3CA*, além de uma alta prevalência de hipermetilação no DNA; ii) tumores com instabilidade de microssatélites (**MSI**), caracterizados por apresentarem elevadas taxas mutacionais, podendo ser classificados como de nível baixo (**MSI-L**) ou de nível alto (**MSI-H**), sendo estes últimos normalmente associados ao **AdG** intestinal e com melhor prognóstico; iii) tumores genomicamente estáveis (**GS**, *Genomically Stable*), caracterizados por possuírem uma frequência relativamente baixa de mutações no gene *TP53*, sendo mais comuns em pacientes jovens; e, por fim, iv) tumores cromossomicamente instáveis (**CIN**, *Chromosomal Instability*), caracterizados por apresentarem altas taxas mutacionais nos genes *TP53*, *ARID1A*, *KRAS*, *PIK3CA*, *RNF43*, *ERBB2* e *APC*, além de fosforilação significativamente elevada em Receptores do Fator de Crescimento Epidérmico (**EGFR**, *Epidermal Growth Factor Receptor*), sendo o subtipo molecular mais frequente no **CG** (BASS et al. 2014).

### 1.3.3 Instabilidade por Microssatélites (MSI)

Microssatélites são sequências curtas, difusas e repetitivas de DNA, distribuídas de forma aleatória no genoma humano (LEUNG et al. 2000). A instabilidade de microssatélite (MSI) é um fenótipo hipermutado associado à perda da atividade do mecanismo de reparo de DNA (MMR, *Mismatch Repair*) (BOLAND e GOEL 2010). A inativação de genes da via *MMR* leva a um acúmulo de erros durante a replicação, como inserções ou deleções de bases em regiões de microssatélites, que se acumulam em alta frequência uma vez que não podem ser corrigidos (ARZIMANOGLOU et al. 1998; FANG et al. 2012).

Em 1997, o Instituto Nacional do Câncer americano (*NCI, National Cancer Institute*) validou e recomendou um painel com cinco marcadores para a classificação do perfil MSI, contendo repetições mono (BAT-25 e BAT-26) e dinucleotídicas (D5S346, D2S123 e D17S250) (BOLAND et al. 1998). O critério para classificação com base nestes marcadores consiste no fato de se dois ou mais des-

tes demonstrarem instabilidade no número de repetições, os genótipos agrupam-se em alta frequência (MSI-H). Quando apenas um dos marcadores demonstra instabilidade, os genótipos agrupam-se em baixa frequência (MSI-L). No caso da ausência de instabilidade nos marcadores, estes microssatélites classificam-se como estáveis (MSS) (BOLAND et al. 1998; SHEMIRANI et al. 2011).

De acordo com as diretrizes de Bethesda (UMAR et al. 2004), as repetições mononucleotídicas demonstraram maior sensibilidade na detecção de MSI quando comparadas às repetições dinucleotídicas. De acordo com este critério, a instabilidade em *loci* de repetições mononucleotídicas define um padrão de MSI-H, enquanto que a instabilidade em apenas loci de repetições dinucleotídicas indicam um padrão de MSI-L (UMAR et al. 2004; BRENNETOT et al. 2005).

Além do método de detecção de MSI através da reação em cadeia da polimerase (PCR), utilizam-se, também, a imuno-histoquímica. Enquanto a PCR analisa a distribuição do comprimento dos microssatélites dentro dos marcadores estabelecidos, a imuno-histoquímica faz a detecção de quatro proteínas associadas à via de *MMR* (MSH2, MSH6, MLH1 e PMS2) (BOLAND et al. 1998). Embora a abordagem laboratorial seja reconhecida como padrão-ouro, seu procedimento de detecção é custoso e seu escopo limitado a apenas um pequeno subgrupo de microssatélites. Por este motivo, ferramentas computacionais vêm sendo desenvolvidas para a detecção de MSI através da análise do exoma total via sequenciamento de nova geração (*NGS, Next-Generation Sequencing*) (NIU et al. 2014; SALIPANTE et al. 2014; HUANG et al. 2015; KAUTTO et al. 2017). Este tipo de abordagem permite uma investigação mais abrangente do perfil de MSI e, possivelmente, será uma aliada nos procedimentos de diagnóstico e tratamento.

#### 1.3.4 MSI e prognóstico

O perfil **MSI** foi descrito pela primeira vez em carcinoma colorretal hereditário não poliposo (HNPCC) por AALTONEN et al. (1993), desde então, vem

sendo descrito em diversas neoplasias humanas hereditárias e esporádicas (ARZIMANOGLU et al. 1998). Este perfil molecular foi associado à carcinogênese e é utilizado como marcador prognóstico em alguns tipos tumorais. Sua relevância clínica é bem estabelecida em câncer colorretal, onde pacientes com tumores MSI-H apresentam melhor prognóstico em comparação à pacientes com tumores MSS (BUCKOWITZ et al. 2005; BENATTI et al. 2005).

Em câncer gástrico, estudos sugerem que tumores **MSI** encontram-se associados ao subtipo histológico intestinal de Lauren e com uma taxa reduzida de metástase linfonodal (LEE et al. 2002; FALCHETTI et al. 2008). No entanto, a associação entre o prognóstico e este perfil molecular ainda permanece inconclusiva (OKI et al. 2009; SEO et al. 2009; FANG et al. 2012; CHOI et al. 2014; ZHU et al. 2015). Estudos apontam MSI-H como marcador de bom prognóstico (SANTOS et al. 1996; BEGHELLI et al. 2006; CHIAVARALLI et al. 2006; FALCHETTI et al. 2008; CORSO et al. 2009; KIM et al. 2011), enquanto outros demonstram resultados conflitantes, principalmente devido ao número limitado de amostras e às diferenças em fatores epidemiológicos considerados nestes estudos (WIRTZ et al. 1998; PEREZ et al. 2004; AN et al. 2005; OKI et al. 2009). Portanto, ainda é necessário uma análise ampla e sistematizada para a compreensão do impacto deste perfil molecular no prognóstico do câncer gástrico.

### 1.3.5 MSI e Imunoterapia

O sistema imunológico envolve um equilíbrio complexo, no qual o processo de identificação e erradicação de antígenos ocorre em conjunto com mecanismos necessários para o controle da intensidade desta resposta imune. Células T CD8+, efectoras ou citotóxicas, são capazes de reconhecer antígenos próprios e não-próprios, ligados a complexos de histocompatibilidade de classe I, que são expressos em APCs. A regulação da resposta mediada por células T pelas células apresentadoras de antígenos (APCs) é parte fundamental na vigilância e elimi-



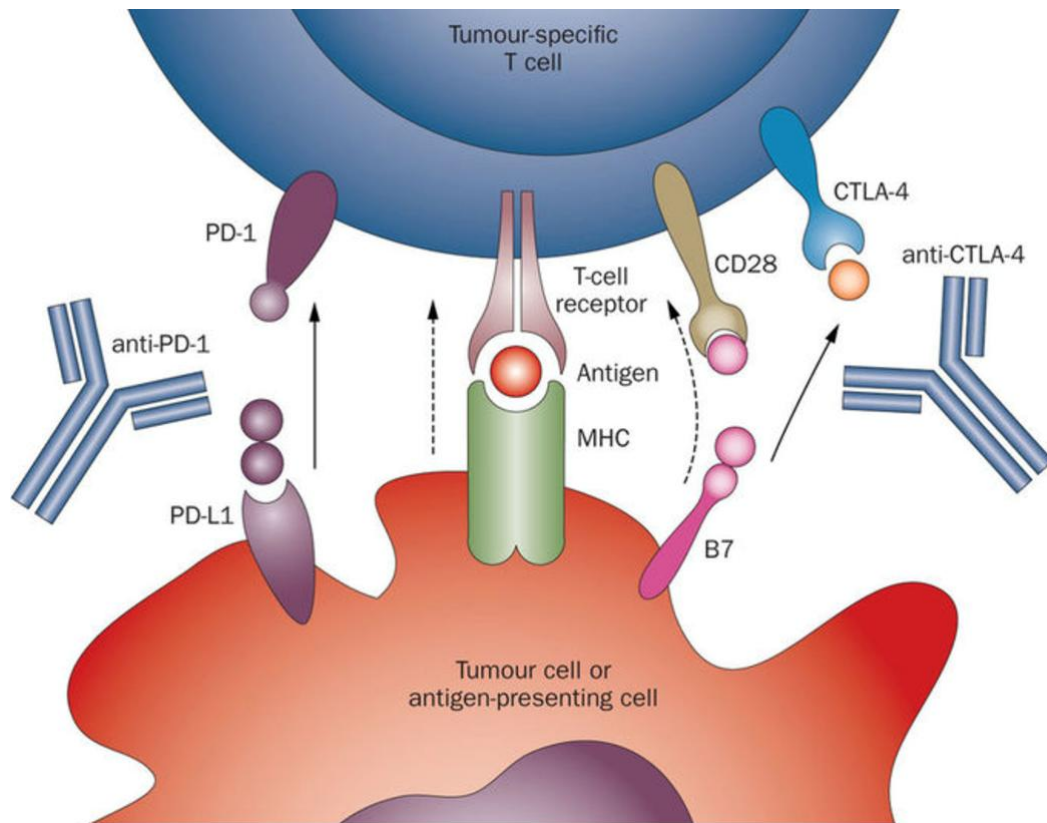
nação de células cancerosas (DIAMOND et al. 2011; FUERTES et al. 2011).

Devido a processos de mutação genética, as células cancerosas produzem neoantígenos que, por sua vez, não são reconhecidos como antígenos próprios pelo sistema imune do hospedeiro levando, assim, a uma resposta direcionada contra o tumor. Como consequência, ocorre a expansão de células T citotóxicas, capazes de identificar as células apresentadoras de neoantígenos, neste caso as células cancerosas, a fim de eliminá-las (DINE et al. 2017).

Em 2013, CHEN e MELLMAN (2013) desenvolveram um modelo para promover uma melhor compreensão dos mecanismos celulares envolvidos na imunidade antitumoral. Este modelo descreve os principais eventos ligados sequencialmente, que devem ocorrer de forma recursiva para montar uma resposta imune eficaz contra a célula tumoral. Além disso, o modelo destaca a importância dos *checkpoints* imunológicos responsáveis por impedir que as células T identifiquem a célula tumoral (CHEN e MELLMAN 2013). Com base nestes achados, estratégias terapêuticas focadas no bloqueio destes processos inibitórios do ciclo de imunidade antitumoral têm ganhado destaque (MELERO et al. 2015).

No entanto, o mecanismo de resposta imune contra uma célula tumoral raramente é eficaz (MELERO et al. 2015). Estudos recentes apontam que vias inibitórias de sinalização, ou “*checkpoints* imunes”, são capazes de impedir a resposta imune contra o tumor (MULLARD 2013). As células tumorais possuem dois mecanismos principais para evitar a ação do sistema imune. O primeiro envolve a via do Antígeno 4 associado ao Linfócito T Citotóxico (CTLA-4), e o segundo associa-se à proteína de morte celular programada 1 (PD-1) e ao seu ligante (PD-L1) (MELLMAN et al. 2011). Células T ativadas expressam PD-1, enquanto que células tumorais expressam PD-L1 em sua superfície (Figura 3) (FREEMAN et al. 2000; OKAZAKI e HONJO 2007).

Uma vez que a interação entre as moléculas PD-1 e PD-L1 resulta na inibição da migração e proliferação das células T, a inserção de um mecanismo capaz de bloquear esta interação resultará no aumento da resposta imune do hospedeiro,



Fonte: DRAKE et al. (2014)

**Figura 3** – Mecanismo de bloqueio de *Checkpoint* imunológico. Esta abordagem à imunoterapia é exemplificada por anticorpos dirigidos contra *CTLA-4*, e por um segundo ponto de checagem maior, mediado pela interação entre *PD-1* nas células T e seu ligante *PD-L1*, em células apresentadoras de antígenos ou células tumorais.

permitindo, assim, a atividade antitumoral (HAMANISHI et al. 2016).

A manipulação dos *checkpoints* imunes revolucionou o tratamento do câncer, sendo aprovada para uso em tumores sólidos e hematológicos (MELLMAN et al. 2011; TOPALIAN et al. 2012). Drogas que atuam como Inibidoras de *checkpoint* imunológico foram aprovadas recentemente por várias agências regulatórias para o tratamento de alguns subtipos tumorais em estágios avançados (MELLMAN et al. 2011).

Estudos demonstraram a eficácia desta estratégia para determinados tipos tumorais, tais como melanoma, carcinoma de pulmão de células não-pequenas e tumores de ovário (HAMANISHI et al. 2015; HAMANISHI et al. 2016). No Brasil, foi aprovado o tratamento com estas drogas para melanoma metastático,

câncer de pulmão de células não pequenas metastático ou localmente avançado, linfoma de Hodgkin refratário, carcinoma urotelial metastático, carcinoma renal metastático e carcinoma gástrico metastático, que expresse PD-L1 no tumor e no infiltrado acima de 1<sup>2</sup> (em uma escala de 0 a 3).

De acordo com BRAHMER et al., alguns pacientes com tumores colorretais MSI-H, com falha na via MMR, obtiveram resposta completa quando submetidos à imunoterapia. Esta resposta encontra-se associada à natureza hipermutada deste subtipo tumoral e à expressão de um grande número de proteínas truncadas originadas de mutações *frameshift*, que, ao atuarem como neoantígenos, desencadeiam uma resposta imune por linfócitos infiltrantes (SCHWITALLE et al. 2008; BRAHMER et al. 2012). Estes achados sugerem que erros em mecanismos de reparo de DNA, através da via *MMR*, além de associarem-se ao desenvolvimento do tumor, podem ser tidos como potencial caminho para seu tratamento.

Devido aos resultados promissores de terapias com base em inibidores de *checkpoints* imunológicos para alguns tipos tumorais e ao longo período de estagnação no tratamento do **CG**, o papel destes agentes neste tumor tornaram-se alvos de investigação (SHEKARIAN et al. 2015). De acordo com a classificação molecular, os subtipos **MSI** e **EBV** são tidos como alvos de interesse para estes ensaios, por apresentarem alta imunogenicidade e, portanto, maior probabilidade de resposta à agentes imunoterápicos (MAGALHAES et al. 2018). Estudos recentes utilizando o bloqueio de *checkpoints* imune, através de anticorpos direcionados para as moléculas CTLA-4, PD-1 e PD-L1, obtiveram efeitos clínicos em textbfCG com aprovação de medicação Pembrolizumabe para **CG** em estadiamento avançado. Além do estadiamento, estes tumores deveriam expressar um CPS score mínimo 1, numa escala crescente de 0-3, que avalia a expressão de PD-L1 tanto no tumor quanto nas células imunes infiltrantes citeFrenel2017.

As terapias baseadas em agentes inibidores de PD-1 e PD-L1 revelaram-se

---

<sup>2</sup> combine proportion score

promissoras e possíveis alternativas para o tratamento do **CG**. No entanto, apesar destes achados já estarem sendo aplicados na prática clínica, é importante considerar que o tumor gástrico é altamente heterogêneo, ressaltando a necessidade de uma estratificação mais acurada de pacientes que irão se beneficiar deste tipo de terapia com o foco em melhora de desfechos e uso racional de recursos (Salem et al. 2018). Nesse contexto, estudos adicionais são necessários para identificar biomarcadores capazes de selecionar pacientes com melhor perfil de resposta para este tipo de abordagem terapêutica.

## 2 OBJETIVOS

### 2.1 OBJETIVO GERAL

Identificar novas fontes adicionais de instabilidade genômica capazes de influenciar na carcinogênese gástrica, além de seu possível papel na resposta a terapias e na sobrevida.

### 2.2 OBJETIVOS ESPECÍFICOS

- Identificar variações somáticas de uma única base (**SNV**, *Single Nucleotide Variation*);
- Detectar assinaturas mutacionais e associá-las a processos responsáveis por gerar a carga mutacional;
- Buscar assinaturas mutacionais características do subtipo molecular **MSI** de **AdGs**;
- Avaliar possíveis associações entre as assinaturas e o prognóstico dos pacientes;
- Avaliar possíveis associações entre as assinaturas e o perfil imunológico dos pacientes;
- Validar os achados em uma coorte proveniente da população brasileira.

## 3 METODOLOGIA

### 3.1 DADOS MOLECULARES E CLÍNICOS

Dados clínicos e moleculares foram obtidos de três bancos de dados principais: i) The Cancer Genome Atlas (TCGA); ii) cBioPortal; iii) e The Cancer Genome Consortium (ICGC). Ao total, 756 amostras não redundantes de **AdGs** foram avaliadas, sendo 438 amostras do TCGA, 196 do cBioPortal e 122 do ICGC. A data de acesso das amostras públicas foi em 20 de outubro de 2016. A validação dos nossos achados foi realizada em 33 amostras de pacientes com **AdG**. Esses pacientes foram recrutados no A.C.Camargo e pertencem à coorte do projeto temático da FAPESP (14/26897-0), denominado *Epidemiologia e Genômica de Adenocarcinomas Gástricos no Brasil* (GE4GAC group et al. 2017).

A seleção das características clínicas relacionadas com prognóstico na coorte pública foi realizada em conjunto com membros do corpo clínico do Hospital A.C.Camargo, sendo as seguintes consideradas mais relevantes: i) subtipo molecular (CIN, EBV, GS e MSI); ii) comprometimento linfonodal (N0, N1, N2 e N3); iii) e recidiva (anotados como WITH TUMOR - WT e TUMOR FREE - TF).

### 3.2 IDENTIFICAÇÃO DE ASSINATURAS MUTACIONAIS

Um total de 239,786 variações somáticas de uma única base (SNV, do inglês *SNV*, *Single Nucleotide Variation*) foram recuperadas das bases TCGA, cBioPortal e ICGC. Os *SNVs* foram detectados previamente pelo *pipeline* do **TCGA** utilizando o *software* *mutect2* (CIBULSKIS et al. 2013).

Cada *SNV* foi mapeado em sequências de trinucleotídeos, considerando o contexto da base 5' e 3' adjacente ao nucleotídeo alterado. Ao final, um total de 96 trincas possíveis foram geradas para cada um dos 756 pacientes, resultando

em uma matriz NxM com dimensão 756x96. Essa matriz considerou as versões **hg19** e **hg38** de acordo com *pipeline* utilizado nos respectivos bancos.

As assinaturas mutacionais foram estimadas pelo *software* *signeR* (RO-SALES et al. 2017). O número ótimo de assinaturas para representar os dados foi baseado no valor médio do Critério de Informação Bayesiano (do inglês *BIC*, *Bayesian Information Criterion*).

A similaridade das assinaturas com aquelas disponíveis no COSMIC <sup>1</sup> foi realizada por meio da distância do cosseno. Nessa métrica, o resultado encontra-se no intervalo entre 0 e 1, onde a distância 0 representa nenhuma similaridade e 1, alta similaridade.

### 3.3 ASSOCIAÇÃO ENTRE AS ASSINATURAS MUTACIONAIS E O PROGNÓSTICO

Para investigar a existência de possíveis associações entre as assinaturas identificadas e o prognóstico de pacientes com **AdG**, utilizamos o método (**DES-Differential Exposure Score**) disponível no *signeR*. Esse método aplica o teste de Kruskal-Wallis para comparar a contribuição da exposição, ou enriquecimento, das assinaturas entre grupos de pacientes que compartilham uma dada característica clínica.

Para avaliar a uniformidade da distribuição de uma variável clínica dentro dos grupos T e B, aplicou-se o teste de proporção, disponível em R <sup>2</sup>. Através do teste ANOVA, o valor de  $p < 0,05$  foi considerado significativo.

### 3.4 ESTRATIFICAÇÃO DE PACIENTES

Após a identificação das assinaturas mutacionais, foram definidos dois grupos com base no grau de exposição de cada amostra à uma dada assinatura. Os quartis superior ( $\geq 75\%$ ) e inferior ( $\leq 25\%$ ) foram utilizados como *cutoff*

<sup>1</sup> <https://cancer.sanger.ac.uk/cosmic/signatures>

<sup>2</sup> <http://www.R-project.org/>

para estratificar os pacientes nos grupos dos mais enriquecidos (denominado neste estudo como "T", de *Top*) e dos menos enriquecidos (denominado neste estudo como "B", de *Bottom*).

### 3.5 ASSOCIAÇÃO DAS ASSINATURAS MUTACIONAIS E SOBREVIVÊNCIA

Para avaliar o possível impacto das assinaturas identificadas na sobrevivência global dos pacientes, utilizou-se o estimador *Kaplan-Meier*. Para esta análise, foi considerado um *follow-up* de 1500 dias. Concentramos nossa análise nos pacientes sem metástase à distância ao diagnóstico (amostras classificadas como "M1" pelo estadiamento patológico foram excluídas desta análise). Comparamos os pacientes mais enriquecidos (grupo "T") com os menos enriquecidos (grupo B), através do teste de *Log-rank*. O valor de  $p < 0.05$  foi considerado significativo.

### 3.6 ASSOCIAÇÃO DAS ASSINATURAS ÀS MUTAÇÕES EM GENES DAS VIAS DE REPARO DE DNA

Para validar a associação entre as assinaturas mutacionais indicadoras de bom prognóstico e o processo mutacional associado à falha no mecanismo da via de reparo de DNA *MMR*, verificou-se o total de genes associados à via *MMR* com a pelo menos uma mutação somática (RONEN e GLICKMAN 2001; WOOD et al. 2005; LANGE et al. 2011) em pacientes dos grupos T e B. Para esta análise considerou-se apenas mutações anotadas pelo *software* VEP<sup>3</sup> como sendo de alto ou moderado impacto (*High* ou *Moderate*), sendo excluídas as mutações classificadas como de baixo impacto (*Low* ou *Modifier*). Para esta análise, aplicou-se o teste de proporção, disponível em R<sup>4</sup>.

<sup>3</sup> <https://www.ensembl.org/vep>

<sup>4</sup> <http://www.R-project.org/>



### 3.7 EXPOSIÇÃO DA ASSINATURA EM DIFERENTES COORTES

Uma simulação de *Monte Carlo* foi realizada para avaliar a distribuição da contribuição das assinaturas encontradas nos pacientes provenientes das três distintas coortes (TCGA, cBioPortal e ICGC). Para comparação da distribuição, utilizou-se o teste de *Kruskal-Wallis*. O valor de  $p < 0,05$  foi considerado significativo.

### 3.8 ASSINATURAS MUTACIONAIS E COMPOSIÇÃO DO INFILTRADO INFLAMATÓRIO

A avaliação da composição do infiltrado inflamatório foi realizada através de um método analítico de deconvolução que é capaz de estimar a proporção de diferentes tipos celulares com base na expressão gênica. Foi construído um painel de referência contendo a expressão gênica das células imunológicas denominado matriz de assinaturas (*signature matrix*). Para a definição destas assinaturas, os valores de expressão foram baseados em genes marcadores específicos (*marker genes*) e em dados de expressão de células do sistema imune purificadas disponíveis na literatura (*reference datasets*).

A seleção de *marker genes* considera a capacidade destes genes de diferenciar os subtipos celulares imunes entre si, bem como a ausência de expressão destes em células estromais ou tumorais. Dentre as listas de *marker genes* existentes na literatura, foi selecionada a LM22 (NEWMAN et al. 2015), composta por 547 genes, cuja assinatura foi validada para 22 tipos celulares imunes.

A etapa seguinte foi a construção do painel de referência, onde se adotou a estratégia proposta por (LI et al. 2017), que utiliza o *pipeline* TIMER e baseia-se na lista de genes do LM22, no banco de dados do *Human Primary Cell Atlas* (HPCA) (MABBOTT et al. 2013) e na expressão gênica (em *FPKM*, do inglês *Fragments Per Kilobase Million*) do tumor em questão. O banco HPCA possui dados de expressão para vários tipos celulares imunes, sendo que para esse estudo

foram escolhidos 3 tipos capazes de estratificar os linfócitos principais: células **B**, **T CD4+** e **T CD8+**. O algoritmo *lsfit* (ABBAS et al. 2009) contido no pacote *CellMix* (GAUJOIX e SEOIGHE 2013) foi utilizado na etapa de deconvolução de dados de expressão global.

Em conjunto com o Laboratório de Imuno-Oncologia Translacional do A.C. Camargo Cancer Center, foram selecionados genes marcadores (GZMA, GZMB, IFNG, PRF1, TNF e CD274/PD-L1) associados com o perfil de linfócitos citotóxicos a fim de caracterizar esse perfil nos grupos B e T.

O teste de correlação de *Pearson* foi aplicado para avaliar as relações entre os genes de linfócitos citotóxicos e a presença de células T CD8+ em pacientes dos grupos T e B. Para esta análise, foi aplicado o Teste de Wilcoxon e o valor de  $p < 0.05$  considerado significativo.

### 3.9 AVALIAÇÃO DO STATUS MSI

A predição do *status MSI* nas amostras de **AdG** foi realizada utilizando-se o *software* MSIsseq (HUANG et al. 2015). Ao contrário dos testes laboratoriais que avaliam o comprimento de poucas regiões repetitivas, abordagens utilizando o sequenciamento de nova geração são capazes de analisar um número superior de regiões, oferecendo uma visão mais completa da frequência de mutações somáticas. A maior parte dos *softwares* (NIU et al. 2014; SALIPANTE et al. 2014; BONNEVILLE et al. 2017) existentes na literatura determinam o *status MSI* a partir de arquivos brutos provenientes do sequenciamento, que nem sempre se encontram disponíveis e demandam alta capacidade de processamento e tempo. Por este motivo, MSIsseq é considerado mais rápido e simples que os demais, uma vez que é capaz de determinar o *status MSI* a partir de um arquivo com as mutações no formato MAF (MAF, do inglês *Mutation Annotation Format*). Este *software* define o *status MSI* das amostras com base nos seguintes critérios: i) SNVs e microinserções/deleções em regiões de repetições de mono-

nucleotídeos (comprimento  $\geq 5$ ) e microssatélites (di-, tri- ou tetra-nucleotídeos anotados como “ *simple repeats*” no banco de dados UCSC <sup>5</sup>; ii) comprimento total da sequência-alvo para obter a contagem de mutações por megabase; iii) tipo de câncer como variável de entrada possível, uma vez que a assinatura mutacional de **MSI** pode variar de acordo com subtipo tumoral. O classificador foi baseado em plataformas de aprendizado de máquina fornecidas pelo pacote, disponível em R, RWeKa <sup>6</sup>.

### 3.10 ANÁLISE DA EXPRESSÃO DE *PD-L1* NO INFILTRADO INFLAMATÓRIO E NO TUMOR

A avaliação da pureza do tumor foi realizada utilizando *scores* imunológicos gerados pelo método *ESTIMATE* (do inglês, *Estimation of STromal and Immune cells in MAlignant Tumours using Expression data*)(YOSHIHARA et al. 2013). Este método <sup>7</sup> considerou um conjunto assinaturas de expressão gênica para inferir a fração de células estromais e imunológicas em amostras de 11 tipos tumorais, incluindo **AdG**, disponíveis no *TCGA*. Essa análise permitiu avaliar se a expressão gênica do gene *PD-L1* teve como origem as células do tumor. Adicionalmente, para verificar se a concentração de infiltrado inflamatório difere significativamente entre os grupos T e B, foi aplicado o teste de Wilcoxon e o valor de  $p < 0,05$  foi considerado significativo.

### 3.11 ANÁLISE DA COORTE DE VALIDAÇÃO

Uma coorte composta por 33 amostras provenientes de pacientes tratados no A. C. Camargo Cancer Center foi utilizada para validação do método de detecção de assinaturas utilizado neste estudo (ver Métodos 3.2). Analisou-se, também, a associação entre as assinaturas validadas e a sobrevida (ver Métodos

<sup>5</sup> <http://hgdownload.cse.ucsc.edu/goldenpath/hg19/database/>

<sup>6</sup> <http://cran.r-project.org/web/packages/RWeKa/>, version 3.7.2

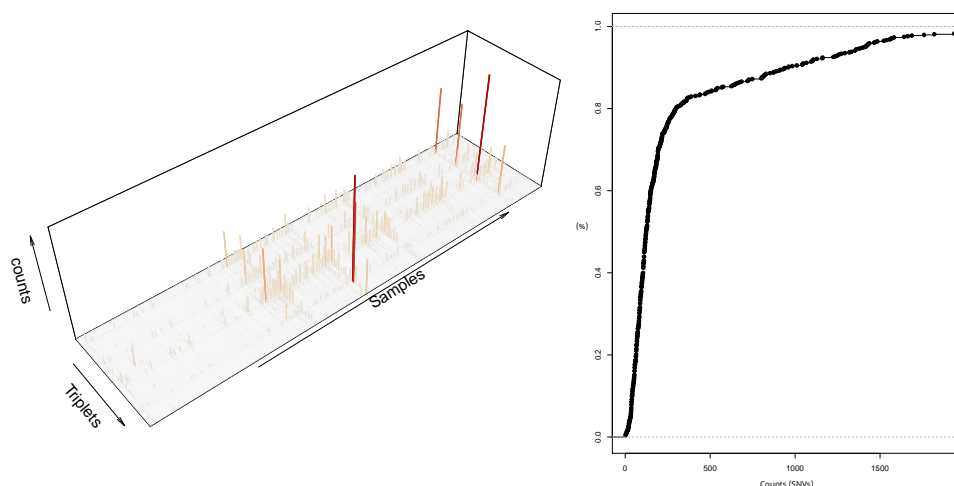
<sup>7</sup> <https://sourceforge.net/projects/estimateproject/>

3.5). E, devido a importância prognóstica, classificaram-se as amostras desta coorte de acordo com o status **MSI**, estimados pelo software MSIseq (ver Métodos 3.9).

## 4 RESULTADOS E DISCUSSÃO

### 4.1 ANÁLISE DO PERFIL MUTACIONAL DAS AMOSTRAS

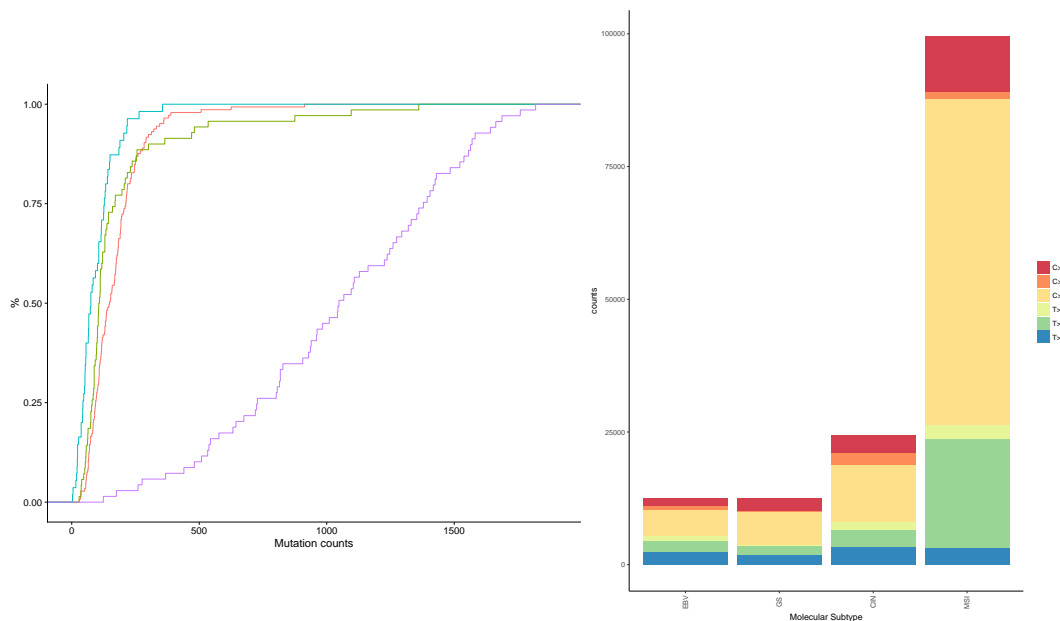
Uma distribuição uniforme do total de *SNVs* é observada na maioria dos tumores que compõem a coorte de 756 amostras de **AdG** (Figura 4), com exceção de um pequeno conjunto de amostras, que demonstrou um padrão mutacional elevado (Figura 4-esq). Ao se analisar a frequência mutacional dentro desta coorte, verificou-se que aproximadamente 80% dos pacientes apresentam até 400 mutações (Figura 4-dir).



**Figura 4** – Matriz de *counts* e distribuição. Representação da matriz de *counts* considerando as 756 amostras de **AdG** (esq.). Gráfico de distribuição acumulada de *SNV*'s nas 756 amostras de **AdG** (dir.).

Ao analisarmos o perfil mutacional dos subtipos moleculares (EBV, MSI, CIN e GS) de **AdG**, identificamos uma carga mutacional superior em tumores classificados como **MSI**, onde aproximadamente 80% das amostras apresentou 1500 mutações ou mais (Figura 5-esq). Adicionalmente, notamos maior frequência de mutações no contexto T>C neste subtipo (Figura 5-dir). Esses resultados

são esperados, uma vez que o perfil **MSI** caracteriza-se por possuir taxas mutacionais elevadas (BASS et al. 2014). No entanto, alguns estudos como de (Salem et al. 2018) relatam que, apesar de **MSI** ser o principal impulsionador da alta carga mutacional (*mutational burden*), este perfil pode ser variável entre amostras de **CG**, sugerindo a necessidade de mais estudos clínicos para avaliar a integração destes perfis a outros fatores moleculares.



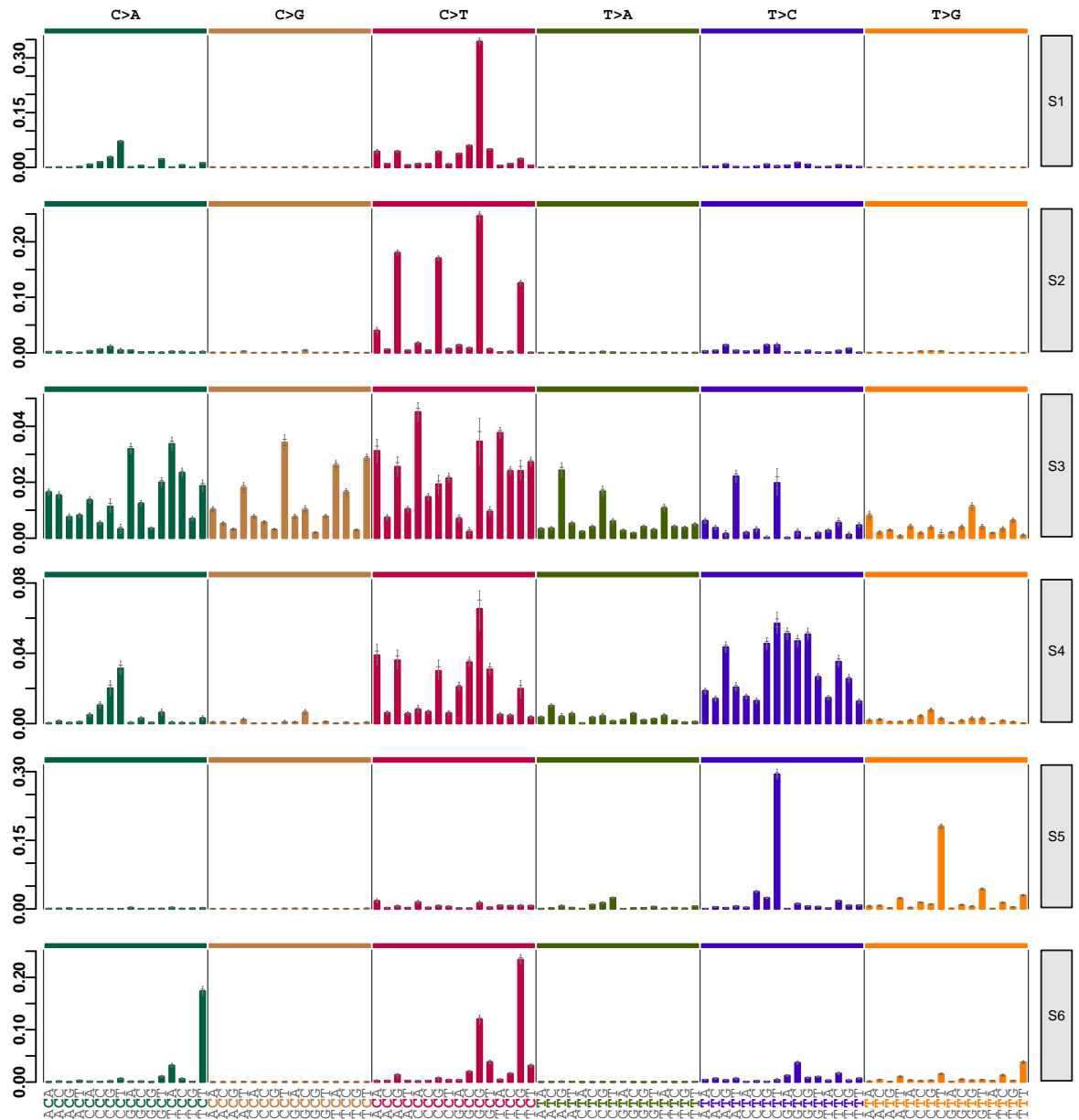
**Figura 5** – Distribuição acumulada e trocas. Distribuição acumulada (esq.) e por troca (dir.) das mutações nos subtipos moleculares (EBV, MSI, GS e CIN).

## 4.2 IDENTIFICAÇÃO DAS ASSINATURAS MUTACIONAIS

Através do *signeR*, seis assinaturas mutacionais foram estimadas no conjunto de 756 amostras de **AdG** (Figura 6). A distância do cosseno foi a métrica empregada para avaliar a similaridade com as assinaturas presentes no banco de dados *COSMIC*<sup>1</sup> (Figura 7).

Esta análise permitiu a identificação dos processos etiológicos que estão ativos em nossa coorte em comparação aos processos já conhecidos. São estes: i) a falha no mecanismo de reparo do DNA (Via **MMR**, *Mismatch Repair*) (S1 e S4);

<sup>1</sup> <http://cancer.sanger.ac.uk/cosmic/>



**Figura 6** – Conjunto de assinaturas mutacionais. Seis assinaturas identificadas pelo método *signeR* nas 756 amostras de Adenocarcinoma gástrico.

ii) a desaminação espontânea da 5-metilcitosina (S2); iii) a falha no mecanismo de reparo de DNA de fita dupla por recombinação homóloga, com forte associação à mutações somáticas e germinativas nos genes *BRCA1* e *BRCA2*; iv) e atividade alterada do gene *POLE*, normalmente associado à amostras hipermutadas (S6). Apenas uma assinatura (S5) possui etiologia desconhecida, embora tenha sido observada em outros subtipos tumorais.

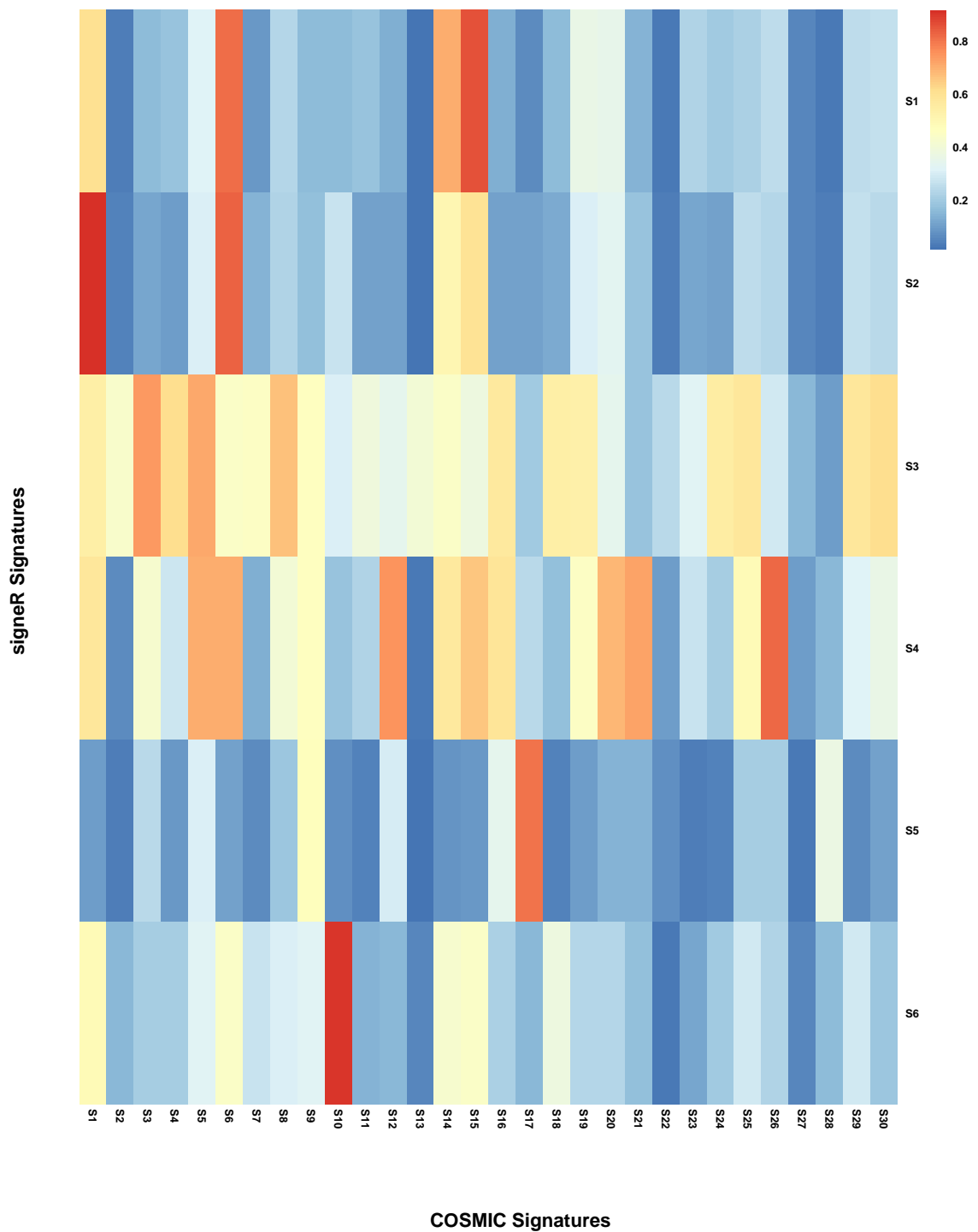
A Figura 7 destaca a similaridade das assinaturas encontradas no presente estudo com aquelas já conhecidas. Adicionalmente, é intuitivo observar que diversos processos mutacionais estão ativos em **AdG**.

### 4.3 ASSOCIAÇÃO DAS ASSINATURAS AO PROGNÓSTICO E SOBREVIDA

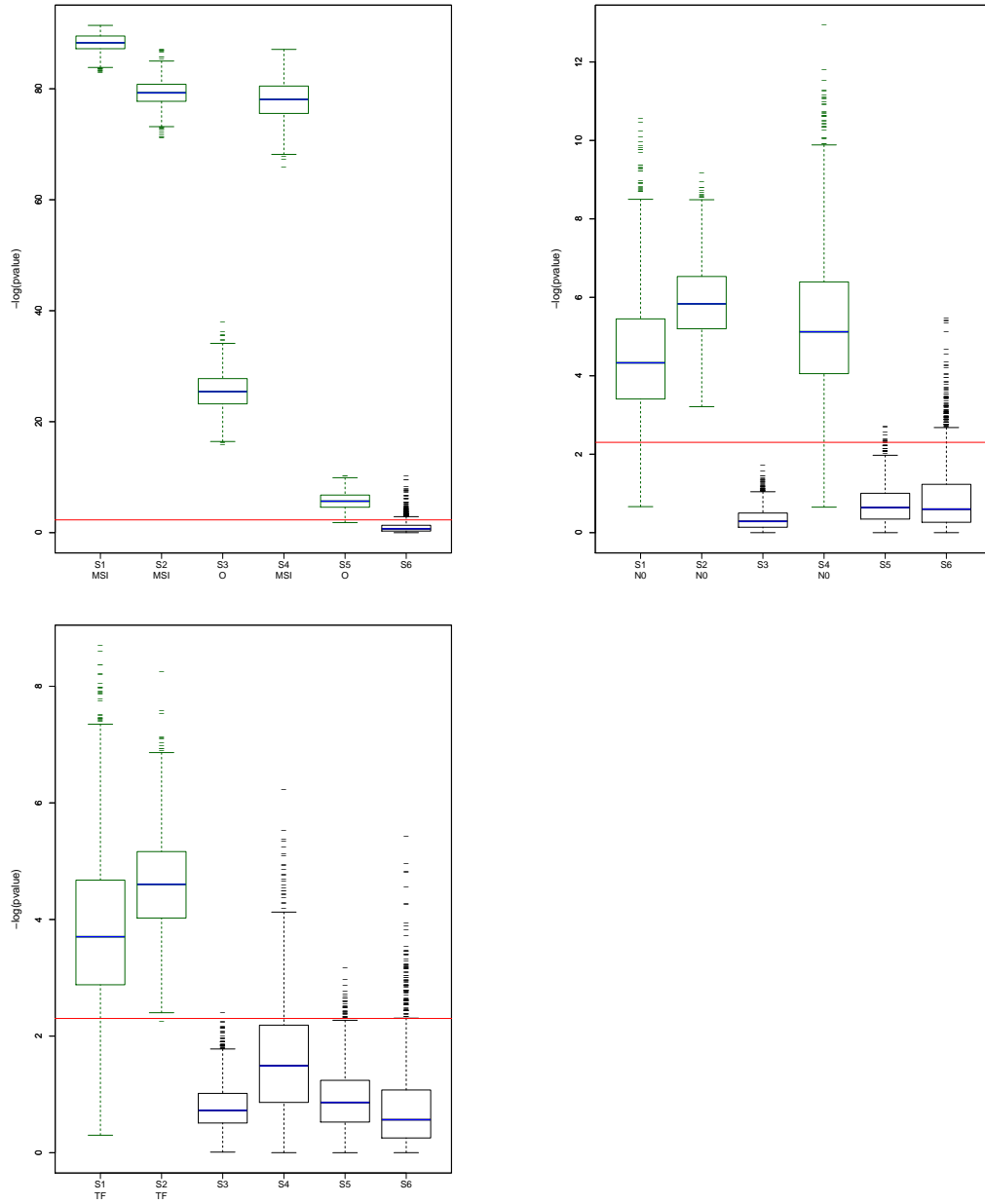
Ao aplicar o método (**DES-Differential Exposure Score**) disponível no *signeR*, identificou-se que tumores **MSI** com ausência de metástase linfonodal (N0) e que não apresentaram recidiva ( *Tumor Free*) são enriquecidos para as Assinaturas S1, S2 e S4 (Figura 8).

Através do estimador *Kaplan-Meier*, observou-se que pacientes do grupo T para estas assinaturas apresentam uma sobrevida melhor ( $p < 0,05$ ) quando comparados àqueles do grupo B (Figura 9). A associação entre **MSI** e as assinaturas *signeR-S1S4* reforçam a importância de genes da via de reparo, uma vez que a inativação dos genes da via MMR resulta em uma alta incidência de pequenas inserções e deleções em sequências de repetições simples, caracterizando o perfil **MSI** (SUPEK e LEHNER 2015).

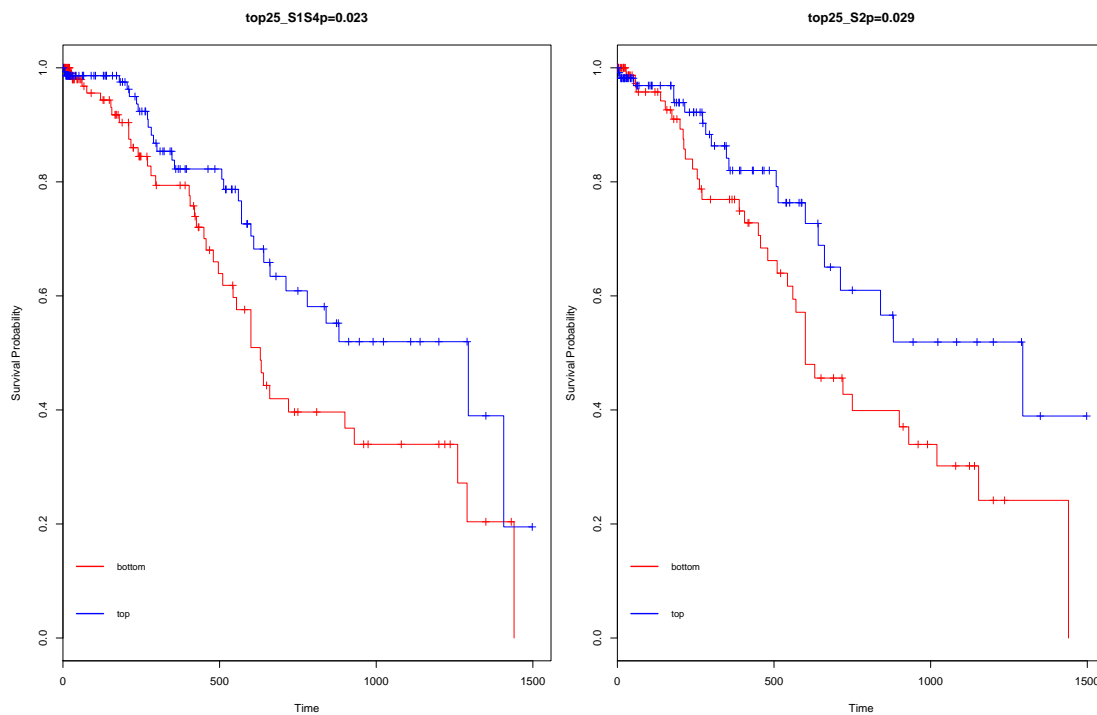




**Figura 7** – *Heatmap* de similaridade das assinaturas com o *COSMIC*. Similaridade, definida pela distância cosseno, entre as assinaturas identificadas pelo *signeR* e as assinaturas presentes no banco de dados *COSMIC*. Os tons mais próximos de vermelho indicam um maior grau de similaridade entre as amostras, enquanto que os tons mais próximos de azul indicam um menor grau de similaridade.



**Figura 8** – Exposição diferencial das assinaturas e perfil clínico. Gráfico de exposição diferencial demonstrando a contribuição das assinaturas associadas às características clínicas com impacto prognóstico: subtipo molecular(sup. esq.), comprometimento linfonodal(sup. dir.) e recidiva (inf. esq.). No eixo x estão as 6 assinaturas e no eixo y a significância de cada assinatura. A mediana das exposições acima da linha vermelha representa a significância ( $p < 0.05$ ) da assinatura à respectiva variável clínica.



**Figura 9** – Curvas de sobrevida global dos pacientes nos grupos **B** e **T**. Curvas de sobrevida global determinadas através do método de Kaplan-Meier dos pacientes de maior e menor contribuição (T vs B) para assinatura mutacional signeR-S1S4 (esq.); Curvas de sobrevida global, determinadas através do método de Kaplan-Meier dos pacientes de maior e menor contribuição (T vs B) para assinatura mutacional signeR-S2. (dir.)

#### 4.4 CARACTERIZAÇÃO CLÍNICA DOS PACIENTES DOS GRUPOS T E B PARA ASSINATURAS MARCADORAS DE BOM PROGNÓSTICO

Pacientes do grupo **T** apresentaram, em sua maioria, tumores classificados molecularmente como **MSI** e, histologicamente, como intestinais, possuindo localização preferencial em antro-Distal (Tabela I). Adicionalmente, através desta caracterização, observa-se que o perfil de melhor sobrevida ocorre de forma independente do estadiamento patológico ("I", Inicial e "A", Avançado) nos grupos T e B (Tabela I e II), uma vez que existe uma proporção similar de pacientes com estadiamentos distintos nos grupos T e B ( $p > 0,05$ ).

Em conjunto, estes dados sugerem que, independente do estágio patológico do tumor, os pacientes mais enriquecidos com a assinatura S1S4 possuem características moleculares específicas que contribuem para uma melhor sobrevida.

**Tabela 1** – Distribuição das características clínicas dos pacientes no grupo T (S1S4) por subtipo molecular. I(Inicial - Estágios I e II); A(Avançado - Estágio III); D(Difuso); I(Intestinal); TF(Tumor Free); WT(With Tumor); N0(Ausência de comprometimento linfonodal); N+(Presença de um ou mais linfonodos comprometidos); AD(Antro/Distal); CP(Cárdia/Proximal); FB(Fundo/Corpo).

	Estágio Patológico			Lauren			Recidiva			TNM(N)			Topografia			
	I	A	NA	D	I	NA	TF	WT	NA	N0	N+	NA	AD	CP	FB	NA
CIN	16 (44%)	19 (53%)	1 (3%)	3 (8%)	19 (53%)	14 (39%)	28 (78%)	3 (8%)	5 (14%)	12 (33%)	24 (67%)	0 (0%)	15 (42%)	7 (19%)	13 (36%)	1 (3%)
EBV	2 (22%)	7 (78%)	0 (0%)	4 (44%)	0 (0%)	5 (56%)	5 (56%)	1 (11%)	3 (33%)	3 (33%)	6 (67%)	0 (0%)	1 (11%)	2 (22%)	6 (67%)	0 (0%)
GS	2 (50%)	2 (50%)	0 (0%)	1 (25%)	1 (25%)	2 (50%)	3 (75%)	1 (25%)	0 (0%)	1 (25%)	3 (75%)	0 (0%)	0 (0%)	2 (50%)	2 (50%)	0 (0%)
MSI	36 (53%)	26 (38%)	6 (9%)	7 (10%)	30 (44%)	31 (46%)	46 (68%)	4 (6%)	18 (26%)	27 (40%)	37 (54%)	4 (6%)	37 (55%)	6 (9%)	22 (32%)	3 (4%)
NA	25 (30%)	48 (59%)	9 (11%)	7 (8%)	31 (38%)	44 (54%)	24 (29%)	11 (14%)	47 (57%)	21 (26%)	60 (73%)	1 (1%)	22 (27%)	9 (11%)	12 (15%)	39 (47%)

**Tabela 2** – Distribuição das características clínicas dos pacientes do grupo B (S1S4) por subtipo molecular. I(Inicial - Estágios I e II); A(Avançado - Estágio III); D(Difuso); I(Intestinal); TF(Tumor Free); WT(With Tumor); N0(Ausência de comprometimento linfonodal); N+(Presença de um ou mais linfonodos comprometidos); AD(Antro/Distal); CP(Cárdia/Proximal); FB(Fundo/Corpo).

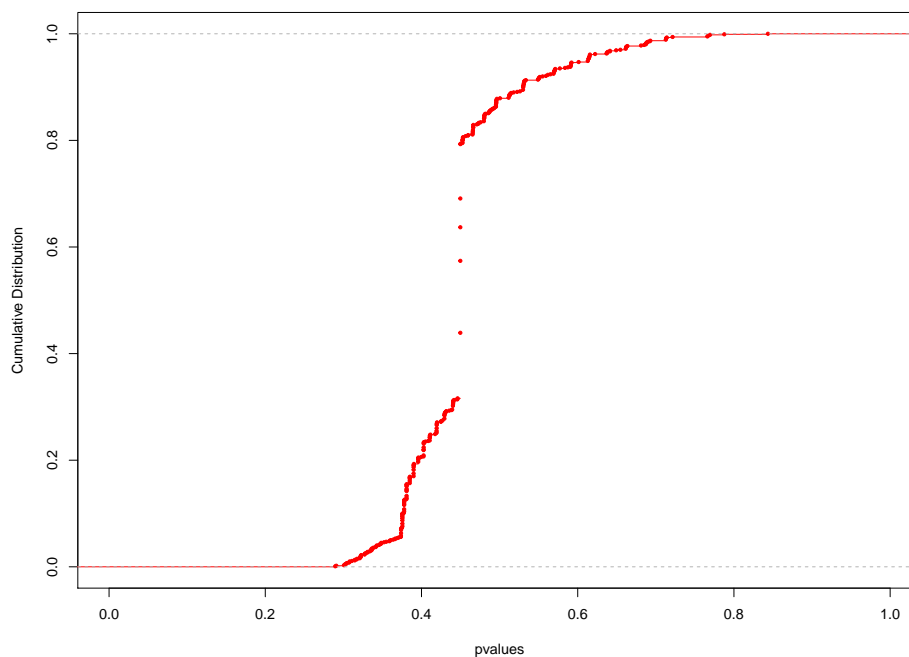
	Estágio Patológico.			Lauren			Recidiva			TNM(N)			Topografia			
	I	A	NA	D	I	NA	TF	WT	NA	N0	N+	NA	AD	CP	FB	NA
CIN	10 (33%)	16 (53%)	4 (14%)	7 (23%)	12 (40%)	11 (37%)	18 (60%)	4 (13%)	8 (27%)	8 (27%)	20 (67%)	2 (7%)	8 (27%)	9 (30%)	12 (40%)	1 (3%)
EBV	3 (25%)	9 (75%)	0 (0%)	2 (17%)	8 (67%)	2 (16%)	8 (66%)	0 (0%)	4 (34%)	1 (8%)	11 (92%)	0 (0%)	4 (33%)	2 (17%)	6 (50%)	0 (0%)
GS	15 (54%)	11 (39%)	2 (7%)	10 (36%)	4 (14%)	14 (50%)	24 (86%)	2 (7%)	2 (7%)	11 (39%)	15 (54%)	2 (7%)	17 (61%)	6 (22%)	4 (14%)	1 (3%)
MSI	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NA	35 (27%)	92 (70%)	3 (3%)	34 (26%)	49 (38%)	47 (36%)	26 (20%)	15 (11%)	89 (69%)	28 (22%)	100 (77%)	2 (1%)	36 (28%)	27 (21%)	30 (23%)	37 (28%)

#### 4.5 MUTAÇÕES EM GENES DE VIAS DE REPARO DE DNA

A via *MMR* é responsável pelo reconhecimento e correção dos erros na biosíntese do DNA causados pela incorporação imprecisa de nucleotídeos durante a replicação. Ao avaliar a associação dos pacientes com mutações em genes da via *MMR* e pacientes com maior exposição da assinatura associada a essa via, foi observado que o grupo T (S1S4) possui um número maior ( $p = 2,655 \times 10^{-13}$ ) de mutações em genes associados à via *MMR*, quando comparados a pacientes do grupo B. Esses resultados destacam a percepção de que a via *MMR* defeituosa leva a um fenótipo hipermutado, uma vez que o grupo T é composto por pacientes com *status* MSI.

## 4.6 CONTRIBUIÇÃO DAS COORTES PÚBLICAS PARA AS ASSINATURAS DE INTERESSE PROGNÓSTICO

Uma vez que o presente projeto avalia três coortes de diferentes bases de dados (TCGA, cBioPortal e ICGC), foi realizada uma simulação de Monte Carlo a fim de se avaliar a uniformidade da contribuição das assinaturas nos pacientes destes 3 bancos. Concluímos que a distribuição da exposição das assinaturas (S1S4) é a mesma, independente das coortes analisadas (Figura 10), uma vez que a hipótese nula não foi rejeitada (Kruskal-Wallis,  $p > 0,05$ ). De particular relevância, é importante destacar a presença de um fenótipo mutador observado em três coortes distintas.

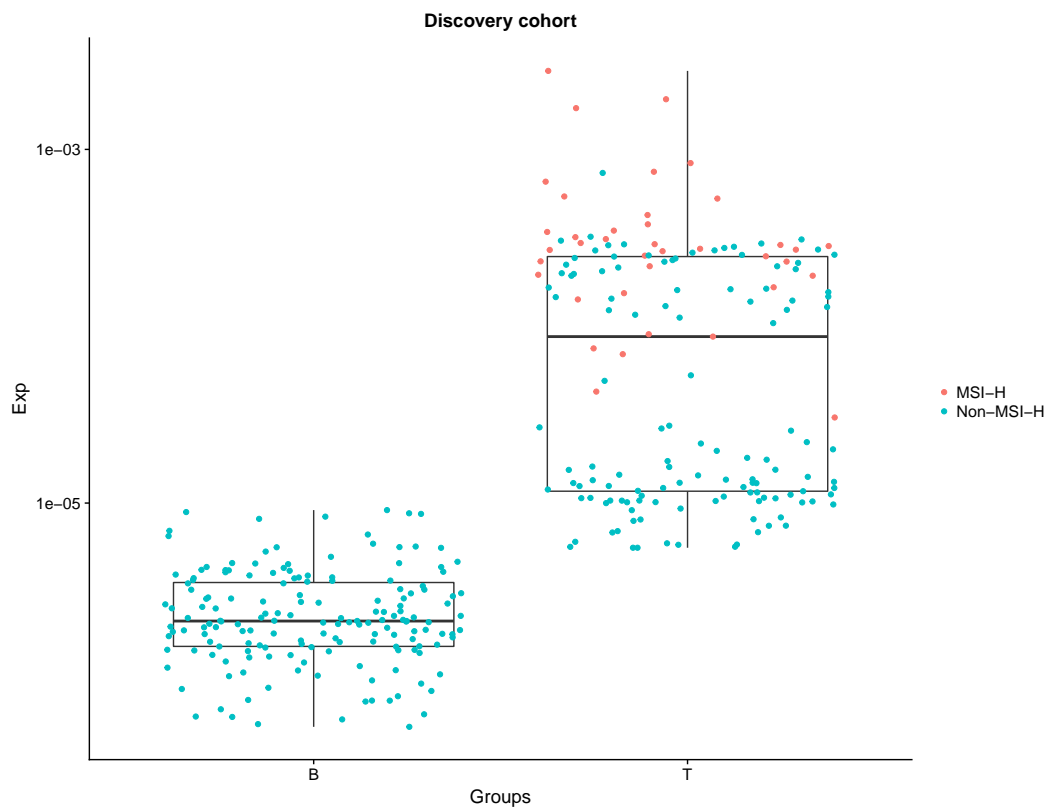


**Figura 10** – Distribuição acumulada da exposição de S1S4 nas três coortes. Gráfico de distribuição acumulada demonstrando a contribuição das assinaturas (S1S4) nos pacientes nas três coortes independentes (TCGA, cBioPortal e ICGC)

### 4.6.1 Classificação da Instabilidade de Microsatélite (MSIseq)

Devido ao fato de apenas 48% das amostras pertencentes à coorte pública possuírem classificação molecular prévia (BASS et al. 2014), utilizamos o

software MSIsseq (ver Métodos 3.9) a fim de estimar o status *MSI* das demais amostras (Figura 11). As amostras com alta exposição à S1S4 pertencentes ao grupo T são, em sua maioria, classificados como MSI-H. E, de maneira complementar, pacientes do grupo B são enriquecidos para amostras com ausência de status MSI-H. Concluímos que a exposição dessa assinatura é capaz de estratificar pacientes MSI-H e com ausência de status MSI-H.



**Figura 11** – Distribuição do *status MSI* nas amostras do grupo T e B para S1S4.

Durante a execução do presente projeto, notamos que abordagens para estimar o status de *MSI* são essenciais para auxiliar na classificação das amostras, principalmente por não limitarem-se a analisar apenas regiões pré-definidas por marcadores. No entanto, os softwares existentes atualmente (NIU et al. 2014; SALIPANTE et al. 2014; BONNEVILLE et al. 2017) oferecem uma classificação binária (MSI-H ou Não MSI-H), onde as amostras com baixo grau de instabilidade (MSI-Low) e as amostras com ausência de instabilidade são agrupadas. Desta forma, o desenvolvimento de um novo classificador, capaz de oferecer um *score*

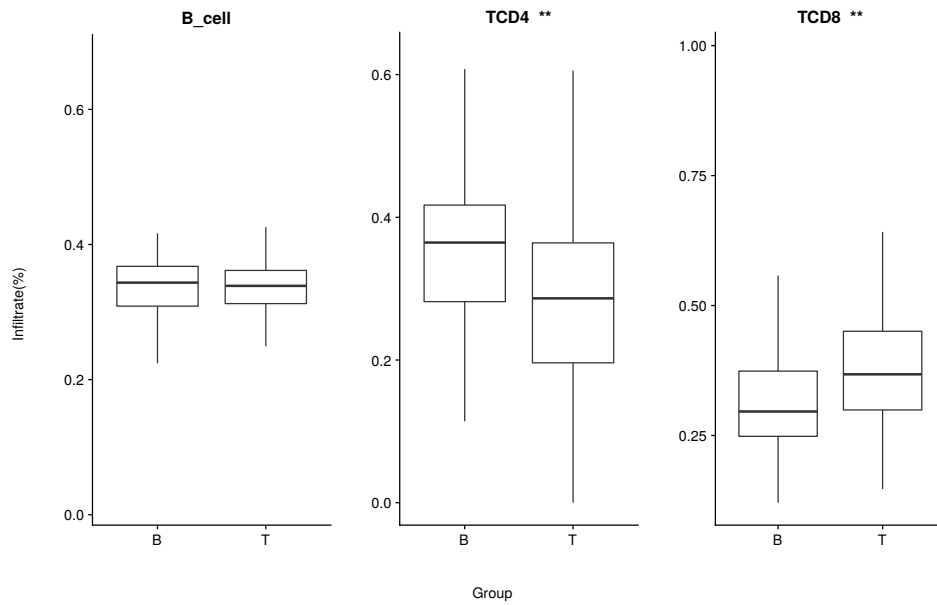
de instabilidade por microssatélites seria de potencial utilidade clínica, uma vez que este perfil é amplamente associado ao prognóstico.

#### 4.7 ASSOCIAÇÃO DAS ASSINATURAS COM PERFIL IMUNOLÓGICO

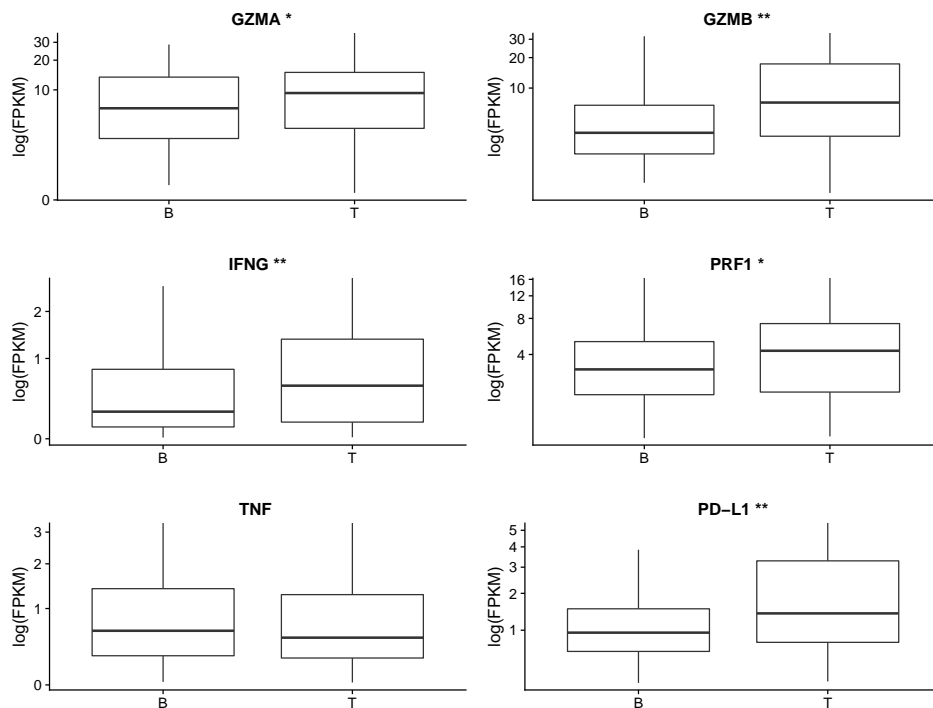
Através da análise do infiltrado inflamatório (ABBAS et al. 2009), observamos que tumores enriquecidos para assinaturas S1S4 (grupo T) apresentaram maior concentração de linfócitos T CD8+, quando comparados aos do grupo B (Figura 12). Atualmente, estudos vêm explorando possíveis associações entre a falha no mecanismo de reparo do DNA e o grau de resposta à imunoterapia (MOUW et al. 2017). Os linfócitos T citotóxicos são considerados componentes essenciais no sistema imune adaptativo, sobretudo por sua alta especificidade no reconhecimento de antígenos e lise das células expressando estes antígenos, incluindo células tumorais (MAHER e DAVIES 2004). Considerando as características exclusivas dos linfócitos T CD8+ como mediadores da imunidade antitumoral, pacientes do grupo T (S1S4), cujos tumores apresentaram maior concentração deste tipo celular, são potenciais candidatos a se beneficiarem de imunoterápicos.

De forma complementar, genes marcadores do perfil imunológico citotóxico (*GZMA*, *GZMB*, *IFNG*, *PRF1* e *TNFA*) e de exaustão (*CD274*/*PD-L1*) também estiveram mais expressos neste mesmo grupo de pacientes. As assinaturas S1 e S4, que compartilham um mesmo processo mutacional associado à falha na via *MMR*, estiveram associadas à expressão de genes envolvidos com citotoxicidade no grupo T quando comparado ao B, com exceção do gene *TNF* (Figura 13). O mesmo ocorre para os genes envolvidos com regulação, que obtiveram sua expressão associada ao grupo B quando comparado ao grupo T, com exceção do gene *IL10* (Figura 14).

Adicionalmente, observamos uma correlação positiva entre os genes marcadores de citotoxicidade e a frequência relativa de células T CD8+ em tumores



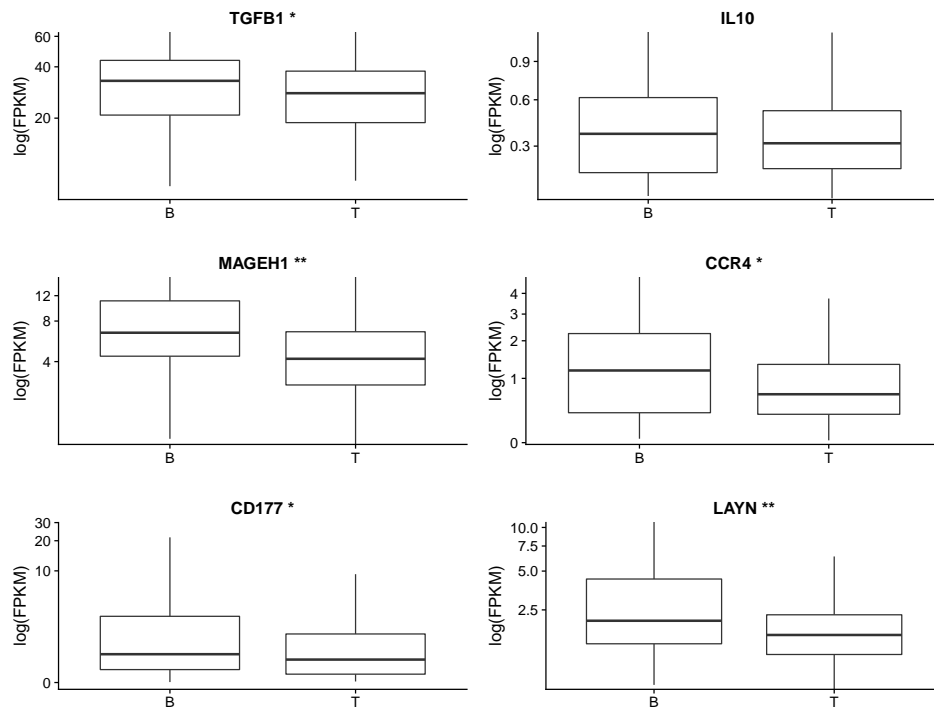
**Figura 12** – Perfil das células B, TCD4 e TCD8 nos grupos **B** e **T**. Perfil das células B, TCD4 e TCD8 nos grupos **B** e **T**. Diferenças significativas ( $p < 0,001$ ) são observadas para as células T CD4+ e T CD8+.



**Figura 13** – Diferença do nível de expressão de genes marcadores de citotoxicidade. Diferença do nível de citotoxicidade e exaustão em tumores de pacientes do grupo T (S1S4) e grupo B (S1S4). Diferenças significativas são destacadas (\* $p < 0,05$ , \*\* $p < 0,01$  - Wilcoxon-test)



de pacientes no grupo T para as assinaturas S1S4 (Figura 14).

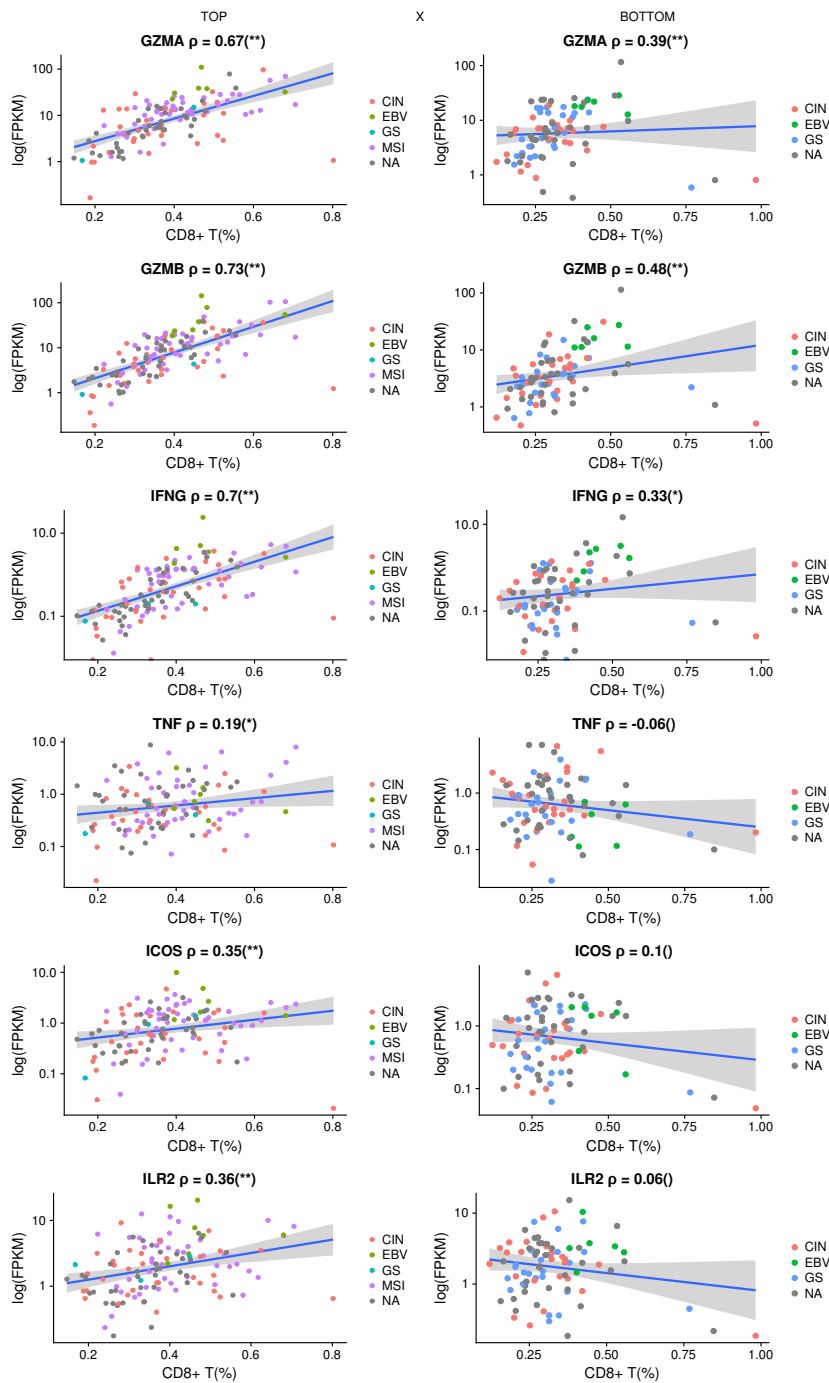


**Figura 14** – Diferença do nível de expressão de genes marcadores de regulação. Diferença observada em tumores de pacientes do grupo T (S1S4) e B (S1S4). Diferenças significativas são destacadas (\* $p < 0,05$ , \*\* $p < 0,01$  - *Wilcoxon-test*)

Ainda dentro deste contexto, observamos uma correlação positiva entre os genes marcadores de citotoxicidade e a frequência relativa de células T CD8+ em tumores de pacientes no grupo T para as assinaturas S1S4 (Figura 15), suportando os achados anteriores.

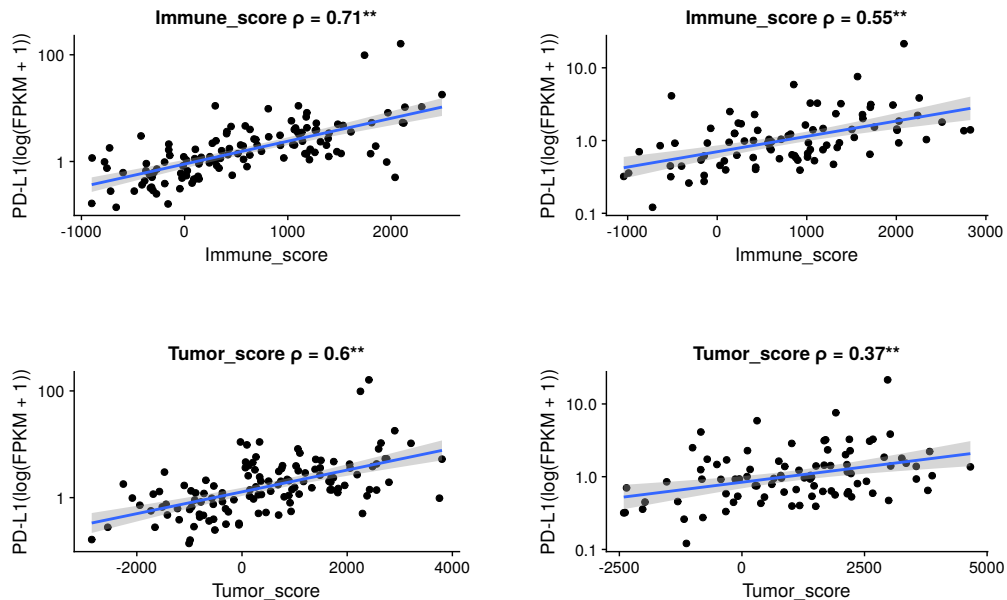
#### 4.8 ANÁLISE DA EXPRESSÃO DE *PD-L1* NO INFILTRADO E NO TUMOR

A fim de avaliar a expressão de *PD-L1* no infiltrado (*Immune score*) e no tumor (*Tumor score*), foi utilizado o *score* gerado pelo método *ESTIMATE* 3.10. Observamos, a existência de correlação entre a expressão de *PD-L1* em ambos infiltrado e tumor para os grupos T e B (S1S4) (Figura 16). Com este resultado, concluímos que o enriquecimento para as assinaturas S1S4 possui potencial poder de estratificação de candidatos à resposta à terapia anti-*PD1/PD-L1*, uma vez



**Figura 15** – Correlação entre expressão de genes marcadores de citotoxicidade e ativação. Correlação do (FPKM) dos genes marcadores de citotoxicidade (*GZMA*, *GZMB*, *IFNG*, *TNF*) e ativação (*ICOS*, *ILR2*) e a frequência relativa do infiltrado de linfócitos T CD8+ em tumores de pacientes do grupo T (S1S4) e B (S1S4). Diferenças significativas são destacadas (\* $p < 0,05$ , \*\* $p < 0,01$  - *Wilcoxon-test*)

que a expressão de *PD-L1* encontra-se não somente no infiltrado inflamatório, mas também nas células tumorais.



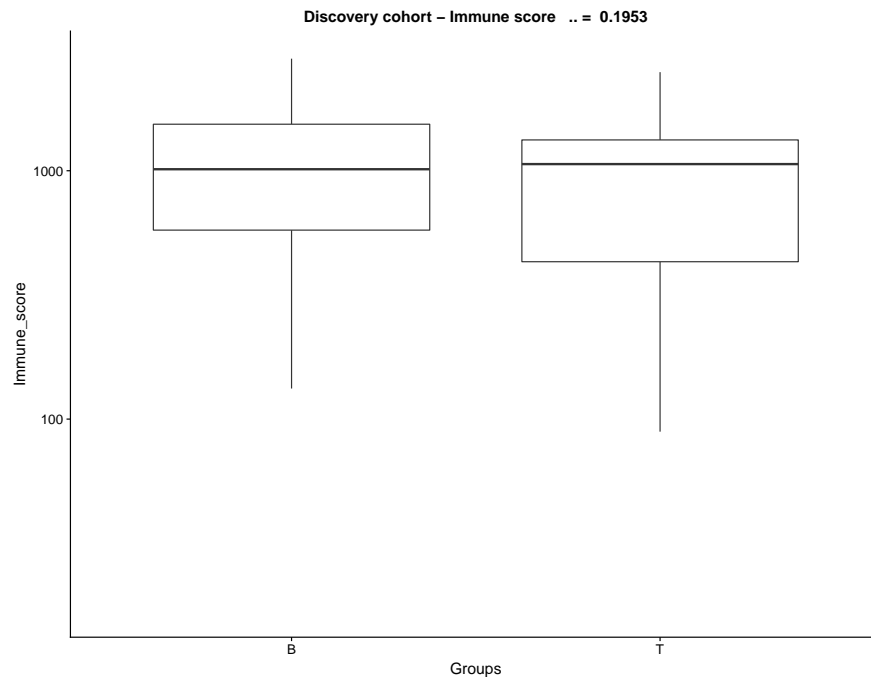
**Figura 16** – Correlação entre a expressão do gene *PD-L1* e das células imunológicas. Correlação entre o (FPKM) do gene *PD-L1* com as células imunológicas (*Immune score*) e tumorais (*Tumor score*) de pacientes do grupo T e B (S1S4). Diferenças significativas são destacadas (\* $p < 0,05$ , \*\* $p < 0,01$  - *Wilcoxon-test*)

A estimação da pureza do tumor foi realizada nos pacientes dos grupos T e B. Essa etapa se fez necessária para avaliar a presença do estroma no tumor a fim de suportar os achados observados na análise infiltrado inflamatório. Conforme representado pela Figura 17, não há diferença significativa entre as medianas entre as amostras dos grupos T e B.

## 4.9 ANÁLISES DA COORTE DE VALIDAÇÃO

### 4.9.1 Assinaturas e Processos Mutacionais

Ao avaliar os processos mutacionais que estão ativos em **AdG** de pacientes da população brasileira, foram detectadas assinaturas em comum (signeR-S1 e signeR-S5) com as coortes previamente estudadas (Figuras 18 e 19). Além das assinaturas que suportam nossos achados prévios, outras assinaturas foram

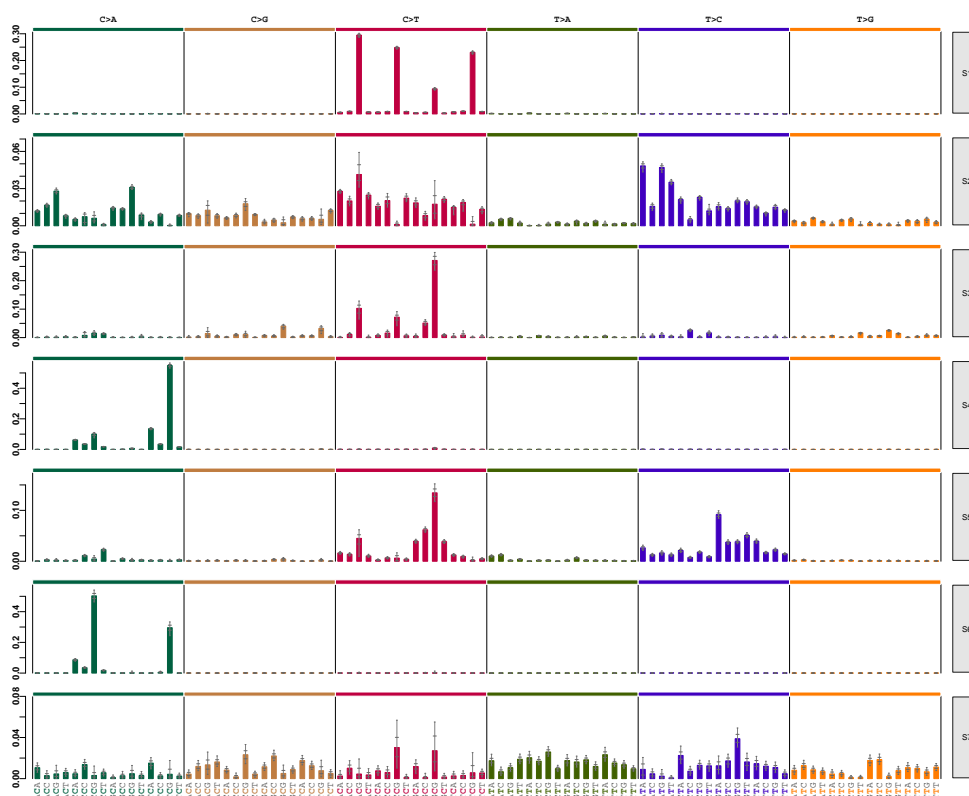


**Figura 17** – Distribuição da composição do infiltrado nas amostras dos pacientes dos grupos ‘B’ e ‘T’ respectivamente.

detectadas: i) *signeR-S2* sem etiologia conhecida, mas detectada em outros tumores; ii) *signeR-S3*, outra assinatura associada à falha no mecanismo de reparo de DNA (*MMR*), sendo descrita em tumores com MSI; iv) *signeR-S4*, *signeR-S6* e *signeR-S7*, sem similaridade com as demais assinaturas já descritas no *COSMIC*, sugerindo um processo específico da coorte analisada. Esses resultados destacam a assinatura associada à via *MMR* em pacientes com **AdG**. Portanto, o diagnóstico da deficiência de *MMR* nos tumores analisados pode apresentar implicações para o manejo desses pacientes.

#### 4.9.2 Análise de sobrevida

Ao contrário da análise anterior, a estratificação dos grupos para a análise de sobrevida na coorte de validação foi realizada mantendo o mesmo critério para o grupo T e, para o segundo grupo, foram incluídos todos os demais pacientes (other). Este critério foi considerado devido ao pequeno número de amostras disponíveis na coorte (N=28) que continham dados de seguimento.

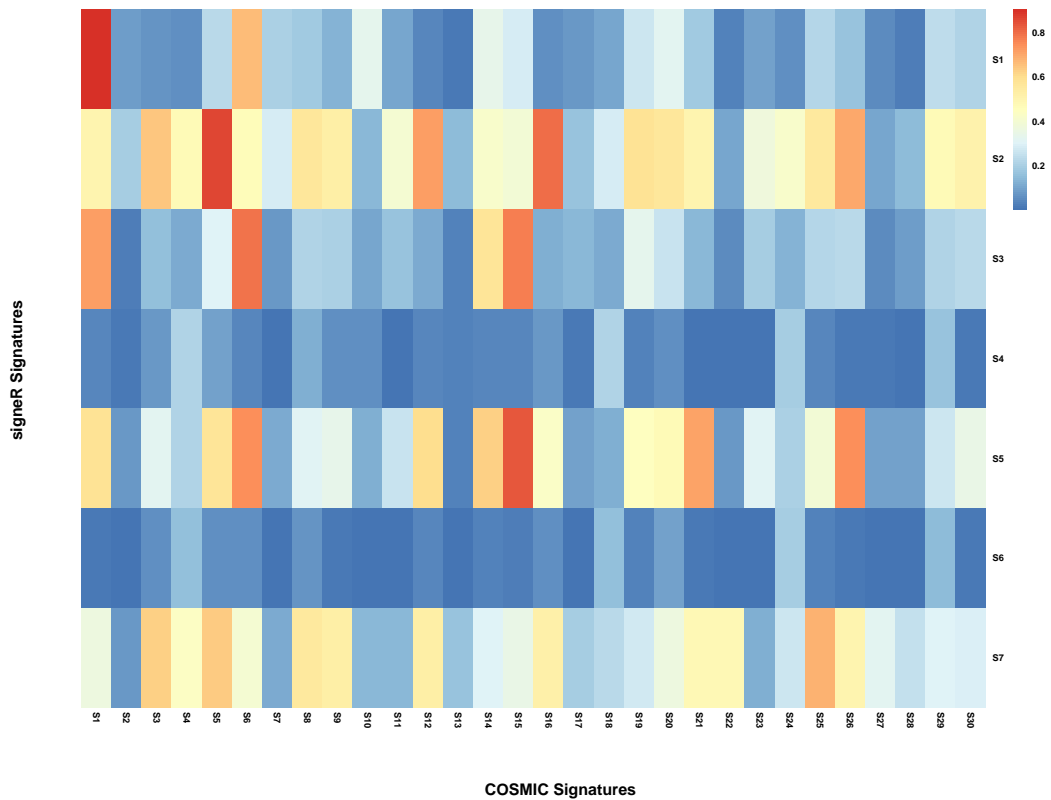


**Figura 18** – Conjunto de assinaturas estimadas pelo método *signeR* nas 33 amostras de Adenocarcinoma Gástrico da coorte prospectiva do A. C. Camargo Cancer Center.

Embora não haja diferença significativa ao comparar os grupos pela função de sobrevivência a partir dos dados de segmento, claramente é observado uma tendência maior de sobrevida dos pacientes com maior exposição (grupo T - *signeR*-S5) quando comparado aos demais pacientes (Figura 20). Esses resultados são concordantes com o que foi observado na análise utilizando um conjunto maior de pacientes de 3 coortes distintas.

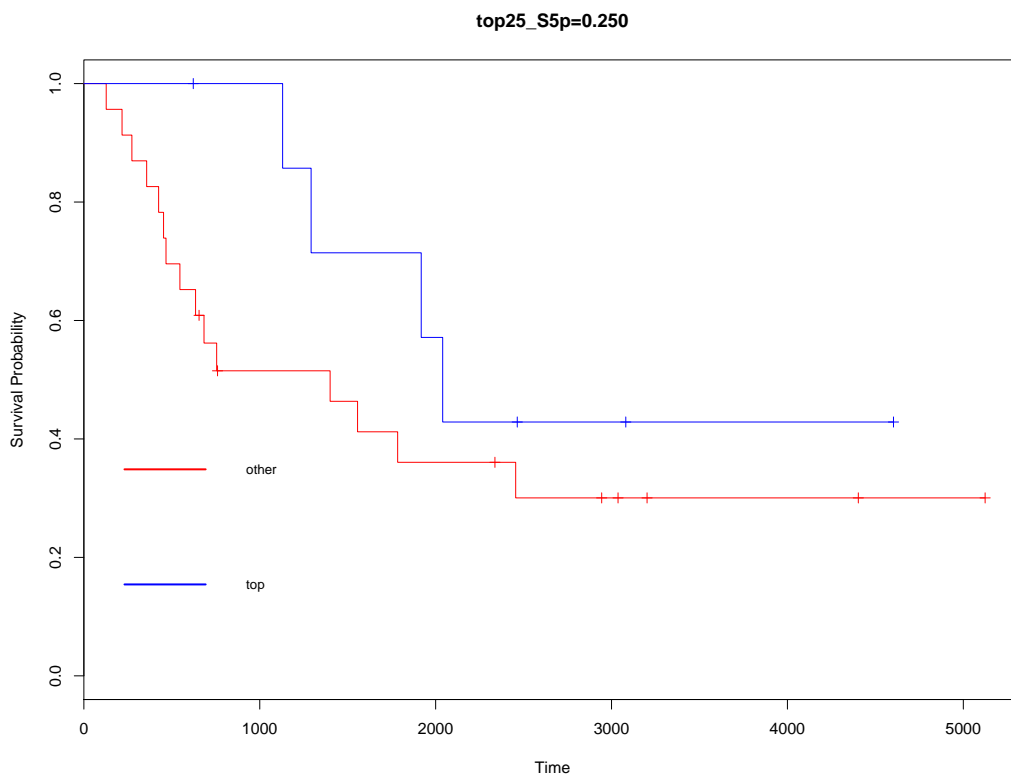
#### 4.9.3 Classificação de Instabilidade por Microssatélite (MSIseq)

O status de *MSI* foi avaliado nos pacientes do grupo T (*signeR*-S5) da coorte do A. C. Camargo Cancer Center. Interessantemente, todos os pacientes desse grupo apresentaram o status MSI-H (Figura 21) estimado pelo software *MSIseq* (HUANG et al. 2015). Este achado é concordante com o que foi observado na coorte pública (Figura 11), uma vez que estas assinaturas compartilham o

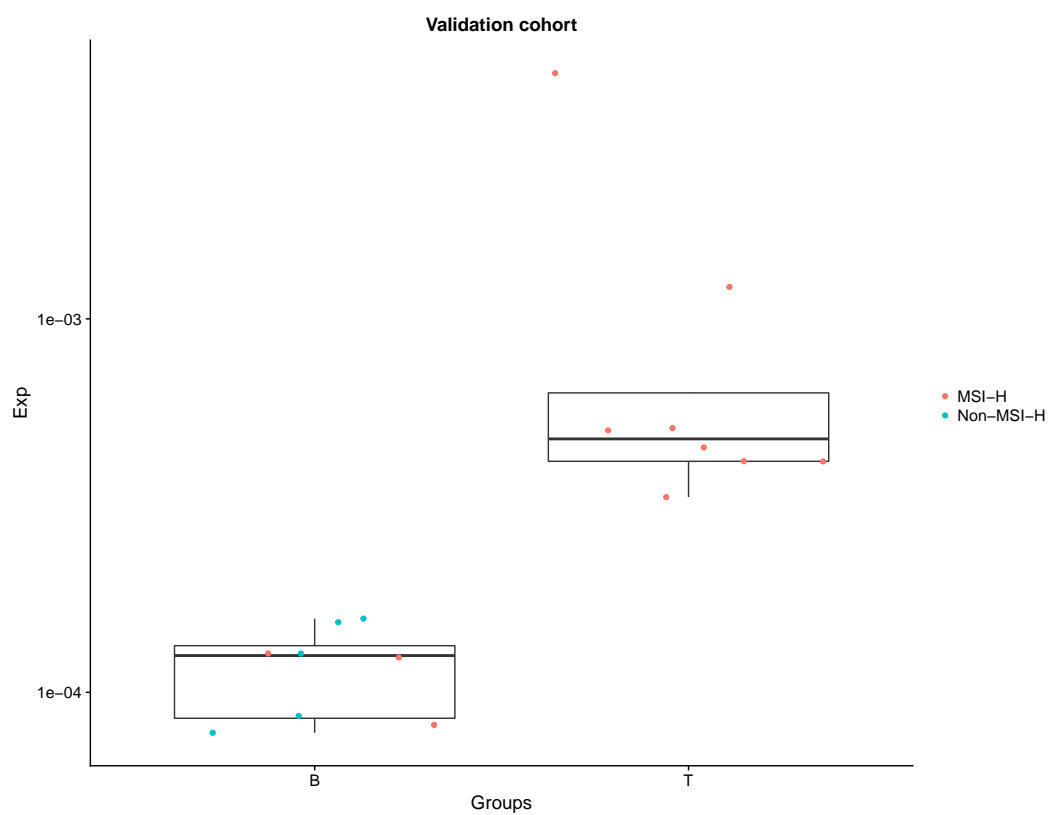


**Figura 19** – Similaridade das assinaturas do *signeR* com as do *COSMIC*. *Heatmap* da similaridade entre as assinaturas identificadas amostras da coorte de **CG** do A. C. Camargo Cancer Center e as assinaturas presentes no banco de dados *COSMIC*. Os tons mais próximos de vermelho indicam um maior grau de similaridade, enquanto que os tons mais próximos de azul indicam um menor grau de similaridade.

mesmo processo etiológico associado à falha na via *MMR*.



**Figura 20** – Curvas de da sobrevida estimadas através método de Kaplan-Meier dos pacientes de maior contribuição e os demais (T vs O) para signeR-S5.



**Figura 21** – Classificação do *status MSI* nos grupos **B** e **T** (signeR-S5).



## 5 CONCLUSÕES

Ao analisar um total de 239,786 mutações somáticas de 756 pacientes com **AdG**, foram detectadas 6 assinaturas mutacionais. Interessantemente, assinaturas associadas à falha no mecanismo de reparo são predominantes no grupo de pacientes pertencente à classe **MSI**, com ausência de metástase linfonodal (N0) e de recidiva. Estes pacientes, por sua vez, apresentam uma melhor sobrevida em comparação àqueles pacientes com menor exposição a esta assinatura. A análise do infiltrado revelou que os tumores de pacientes mais expostos à assinatura associada à falha de MMR possuem maior frequência de células T CD8+ e uma maior expressão de genes associados a um perfil citotóxico e de exaustão, incluindo *PD-L1*. Esses resultados podem ser usados como marcadores para identificar pacientes bons respondedores às terapias anti-*PD1*/anti-*PD-L1* em **CG**. Ao realizar uma análise com um grupo de 33 pacientes tratados no A.C. Camargo Cancer Center, foi detectada a presença da assinatura associada à falha na via *MMR*, validando os achados anteriores. Os pacientes mais enriquecidos para esta assinatura, classificados como MSI-H, também apresentaram melhor sobrevida quando comparados aos pacientes menos enriquecidos. Ao analisar coletivamente todos os resultados, concluímos que a assinatura associada à falha de *MMR* apresenta impacto no prognóstico, além de seu potencial na seleção de pacientes favoráveis à imunoterapia em câncer gástrico.

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

Aaltonen LA, Peltomaki P, Leach FS, et al. Clues to the pathogenesis of familial colorectal cancer. **Science** 1993; 260:812–6.

Abbas AR, Wolslegel K, Seshasayee D, et al. Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. **PLoS ONE** 2009; 4:6098.

Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. **Nature** 2013; 500:415–21.

Alexandrov LB, Stratton MR. Mutational signatures: the patterns of somatic mutations hidden in cancer genomes. **Curr Opin Genet Dev** 2014; 24:52–60.

An C, Choi IS, Yao JC, et al. Prognostic significance of CpG island methylator phenotype and microsatellite instability in gastric carcinoma. **Clin Cancer Res** 2005; 11:656–63.

Aneiros-Fernandez J, Husein-ElAhmed H, Arias-Santiago S, et al. Cutaneous metastasis as first clinical manifestation of signet ring cell gastric carcinoma. **Dermatol Online J** 2010; 16:9.

Arzimanoglou II, Gilbert F, Barber HR. Microsatellite instability in human solid tumors. **Cancer** 1998; 82:1808–20.

Balmain A. Cancer as a complex genetic trait: tumor susceptibility in humans and mouse models. **Cell** 2002; 108:145–52.

Bass AJ, Thorsson V, Shmulevich I, et al. Comprehensive molecular characterization of gastric adenocarcinoma. **Nature** 2014; 513:202–9.

Beghelli S, de Manzoni G, Barbi S, et al. Microsatellite instability in gastric cancer is associated with better prognosis in only stage II cancers. **Surgery** 2006; 139: 347–56.

Benatti P, Gafa R, Barana D, et al. Microsatellite instability and colorectal cancer prognosis. **Clin Cancer Res** 2005; 11:8332–40.

Boland CR, GOEL A. Microsatellite instability in colorectal cancer. **Gastroenterology** 2010; 138:2073–87.

Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. **Cancer Res** 1998; 58:5248–57.

Bonneville R, Krook MA, Kautto EA, et al. Landscape of Microsatellite Instability Across 39 Cancer Types. **JCO Precis Oncol** 2017; 2017:.

Brahmer JR, Tykodi SS, Chow LQ, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. **N Engl J Med** 2012; 366:2455–65.

Brennetot C, Buhard O, Jourdan F, et al. Mononucleotide repeats BAT-26 and BAT-25 accurately detect MSI-H tumors and predict tumor content: implications for population screening. **Int J Cancer** 2005; 113:446–50.

Buckowitz A, Knaebel HP, Benner A, et al. Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. **Br J Cancer** 2005; 92:1746–53.

Cario CL, Witte JS. Orchid: a novel management, annotation and machine learning framework for analyzing cancer mutations. **Bioinformatics** 2018; 34:936–42.

Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. **Immunity** 2013; 39:1–10.

Chiavaralli AM, Feltri M, Bertolini V, et al. Intratumour T cells, their activation status and survival in gastric carcinomas characterised for microsatellite instability and Epstein-Barr virus infection. **Virchows Arch** 2006; 448:344–53.

Choi YY, Bae JM, An JY, et al. Is microsatellite instability a prognostic marker in gastric cancer? A systematic review with meta-analysis. **J Surg Oncol** 2014; 110:129–35.

Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. **Nat Biotechnol** 2013; 31:213–9.

Collins FS, Varmus H. A new initiative on precision medicine. **N Engl J Med** 2015; 372:793–5.

Corso G, Pedrazzani C, Marrelli D, et al. Correlation of microsatellite instability at multiple loci with long-term survival in advanced gastric carcinoma. **Arch Surg** 2009; 144:722–7.

Diamond MS, Kinder M, Matsushita H, et al. Type I interferon is selectively required by dendritic cells for immune rejection of tumors. **J Exp Med** 2011; 208:1989–2003.

Dine J, Gordon R, Shames Y, et al. Immune Checkpoint Inhibitors: An Innovation in Immunotherapy for the Treatment and Management of Patients with Cancer. **Asia Pac J Oncol Nurs** 2017; 4:127–35.

Drake CG, Lipson EJ, Brahmer JR. Breathing new life into immunotherapy: review of melanoma, lung and kidney cancer. **Nat Rev Clin Oncol** 2014; 11: 24–37.

Duell EJ, Travier N, Lujan-Barroso L, et al. Alcohol consumption and gastric cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. **Am J Clin Nutr** 2011; 94:1266–75.

Everatt R, Tamosiunas A, Kuzmickiene I, et al. Alcohol consumption and risk of gastric cancer: a cohort study of men in Kaunas, Lithuania, with up to 30 years follow-up. **BMC Cancer** 2012; 12:475.

Falchetti M, Saieva C, Lupi R, et al. Gastric cancer with high-level microsatellite instability: target gene mutations, clinicopathologic features, and long-term survival. **Hum Pathol** 2008; 39:925–32.

Fang WL, Chang SC, Lan YT, et al. Microsatellite instability is associated with a better prognosis for gastric cancer patients after curative surgery. **World J Surg** 2012; 36:2131–8.

Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. **Int J Cancer** 2015; 136:359–86.

Ferro A, Peleteiro B, Malvezzi M, et al. Worldwide trends in gastric cancer mortality (1980-2011), with predictions to 2015, and incidence by subtype. **Eur J Cancer** 2014; 50:1330–44.

Fischer A, Illingworth CJ, Campbell PJ, et al. EMu: probabilistic inference of mutational processes and their localization in the cancer genome. **Genome Biol** 2013; 14:39.

Freeman GJ, Long AJ, Iwai Y, et al. Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. **J Exp Med** 2000; 192:1027–34.

Fuertes MB, Kacha AK, Kline J, et al. Host type I IFN signals are required for antitumor CD8+ T cell responses through CD8alpha+ dendritic cells. **J Exp Med** 2011; 208:2005–16.

Gaujoix R, Seoighe C. CellMix: a comprehensive toolbox for gene expression deconvolution. **Bioinformatics** 2013; 29:2211–2.

Ge4gac group, Soares FA, Coimbra FJF, et al. Genomics and epidemiology for gastric adenocarcinomas. **Applied Cancer Research** 2017; 37:7.

Gehring JS, Fischer B, Lawrence M, et al. SomaticSignatures: inferring mutational signatures from single-nucleotide variants. **Bioinformatics** 2015; 31:3673–5.

Guastadisegni C, Colafranceschi M, Ottini L, et al. Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data. **Eur J Cancer** 2010; 46:2788–98.

Hamanishi J, Mandai M, Ikeda T, et al. Safety and Antitumor Activity of Anti-PD-1 Antibody, Nivolumab, in Patients With Platinum-Resistant Ovarian Cancer. **J Clin Oncol** 2015; 33:4015–22.

Hamanishi J, Mandai M, Matsumura N, et al. PD-1/PD-L1 blockade in cancer treatment: perspectives and issues. **Int J Clin Oncol** 2016; 21:462–73.

- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. **Cell** 2011; 144:646–74.
- Helleday T, Eshtad S, Nik-Zainal S. Mechanisms underlying mutational signatures in human cancers. **Nat Rev Genet** 2014; 15:585–98.
- Huang MN, McPherson JR, Cutcutache I, et al. MSIseq: Software for Assessing Microsatellite Instability from Catalogs of Somatic Mutations. **Sci Rep** 2015; 5: 13321.
- Hugo W, Zaretsky JM, Sun L, et al. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. **Cell** 2017; 168:542.
- Hwang SW, Lee DH, Lee SH, et al. Preoperative staging of gastric cancer by endoscopic ultrasonography and multidetector-row computed tomography. **J Gastroenterol Hepatol** 2010; 25:512–8.
- Jemal A, Bray F, Center MM, et al. Global cancer statistics. **CA Cancer J Clin** 2011; 61:69–90.
- Kautto EA, Bonneville R, Miya J, et al. Performance evaluation for rapid detection of pan-cancer microsatellite instability with MANTIS. **Oncotarget** 2017; 8: 7452–63.
- Kim H, An JY, Noh SH, et al. High microsatellite instability predicts good prognosis in intestinal-type gastric cancers. **J Gastroenterol Hepatol** 2011; 26: 585–92.
- Kim J, Sun CL, Mailey B, et al. Race and ethnicity correlate with survival in patients with gastric adenocarcinoma. **Ann Oncol** 2010; 21:152–60.
- Lange SS, Takata K, Wood RD. DNA polymerases and cancer. **Nat Rev Cancer** 2011; 11:96–110.
- Lauren P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. **Acta Pathol Microbiol Scand** 1965; 64: 31–49.
- Lee HS, Choi SI, Lee HK, et al. Distinct clinical features and outcomes of gastric cancers with microsatellite instability. **Mod Pathol** 2002; 15:632–40.
- Lee J, Lee AJ, Lee JK, et al. Mutalisk: a web-based somatic MUTation AnaLyIS toolKit for genomic, transcriptional and epigenomic signatures. **Nucleic Acids Res** 2018; 46:102–8.
- Leung WK, Kim JJ, Kim JG, et al. Microsatellite instability in gastric intestinal metaplasia in patients with and without gastric cancer. **Am J Pathol** 2000; 156: 537–43.
- Li T, Fan J, Wang B, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. **Cancer Res** 2017; 77:108–10.

Mabbott NA, Baillie JK, Brown H, et al. An expression atlas of human primary cells: inference of gene function from coexpression networks. **BMC Genomics** 2013; 14:632.

Maconi G, Manes G, Porro GB. Role of symptoms in diagnosis and outcome of gastric cancer. **World J Gastroenterol** 2008; 14:1149–55.

Magalhaes H, Fontes-Sousa M, Machado M. Immunotherapy in Advanced Gastric Cancer: An Overview of the Emerging Strategies. **Can J Gastroenterol Hepatol** 2018; 2018:2732408.

Maher J, Davies ET. Targeting cytotoxic T lymphocytes for cancer immunotherapy. **Br J Cancer** 2004; 91:817–21.

Marrelli D, Roviello F, de Manzoni G, et al. Different patterns of recurrence in gastric cancer depending on Lauren's histological type: longitudinal study. **World J Surg** 2002; 26:1160–5.

Melero I, Berman DM, Aznar MA, et al. Evolving synergistic combinations of targeted immunotherapies to combat cancer. **Nat Rev Cancer** 2015; 15:457–72.

Mellman I, Coukos G, Dranoff G. Cancer immunotherapy comes of age. **Nature** 2011; 480:480–9.

Ministério da Saúde Instituto Nacional de Câncer José Alencar Gomes da Silva. **Estimativa 2014: Incidência de Câncer no Brasil**. INCA, 2014. Disponível em: <<http://www.inca.gov.br/wcm/dncc/2013/apresentacao-estimativa-2014.pdf>>. Acesso em: 16 de abril de 2018.

Ministério da Saúde Instituto Nacional de Câncer José Alencar Gomes da Silva. **Estimativa 2016: Incidência de câncer no Brasil**. INCA, 2016. 122 . Disponível em: <<http://www.inca.gov.br/estimativa/2016/estimativa-2016-v11.pdf>>. Acesso em: 16 de abril de 2018.

Mouw KW, Goldberg MS, Konstantinopoulos PA, et al. DNA Damage and Repair Biomarkers of Immunotherapy Response. **Cancer Discov** 2017; 7:675–93.

Mullard A. New checkpoint inhibitors ride the immunotherapy tsunami. **Nat Rev Drug Discov** 2013; 12:489–92.

Murphy G, Pfeiffer R, Camargo MC, et al. Meta-analysis shows that prevalence of Epstein-Barr virus-positive gastric cancer differs based on sex and anatomic location. **Gastroenterology** 2009; 137:824–33.

Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. **Nat Methods** 2015; 12:453–7.

Nik-Zainal S, Alexandrov LB, Wedge DC, et al. Mutational processes molding the genomes of 21 breast cancers. **Cell** 2012; 149:979–93.

- Niu B, Ye K, Zhang Q, et al. MSI sensor: microsatellite instability detection using paired tumor-normal sequence data. **Bioinformatics** 2014; 30:1015–6.
- Okazaki T, HONJO T. PD-1 and PD-1 ligands: from discovery to clinical application. **Int Immunol** 2007; 19:813–24.
- Oki E, Kakeji Y, Zhao Y, et al. Chemosensitivity and survival in gastric cancer patients with microsatellite instability. **Ann Surg Oncol** 2009; 16:2510–5.
- Perez RO, Jacob CE, D'Ottaviano FL, et al. Microsatellite instability in solitary and sporadic gastric cancer. **Rev Hosp Clin Fac Med Sao Paulo** 2004; 59: 279–85.
- Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. **J Clin Oncol** 2005; 23:609–18.
- Ronen A, Glickman BW. Human DNA repair genes. **Environ Mol Mutagen** 2001; 37:241–83.
- Rosales RA, Drummond RD, Valieris R, et al. signeR: an empirical Bayesian approach to mutational signature discovery. **Bioinformatics** 2017; 33:8–16.
- Sakurai U, Lauwers GY, Vieth M, et al. Gastric high-grade dysplasia can be associated with submucosal invasion: evaluation of its prevalence in a series of 121 endoscopically resected specimens. **Am J Surg Pathol** 2014; 38:1545–50.
- Salem ME, Puccini A, Grothey A, et al. Landscape of Tumor Mutation Load, Mismatch Repair Deficiency, and PD-L1 Expression in a Large Patient Cohort of Gastrointestinal Cancers. **Mol Cancer Res** 2018; 16:805–12.
- Salipante SJ, Scroggins SM, Hampel HL, et al. Microsatellite instability detection by next generation sequencing. **Clin Chem** 2014; 60:1192–9.
- Dos Santos NR, Seruca R, Constancia M, et al. Microsatellite instability at multiple loci in gastric carcinoma: clinicopathologic implications and prognosis. **Gastroenterology** 1996; 110:38–44.
- Schwitalle Y, Kloor M, Eiermann S, et al. Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers. **Gastroenterology** 2008; 134:988–97.
- Seo HM, Chang YS, Joo SH, et al. Clinicopathologic characteristics and outcomes of gastric cancers with the MSI-H phenotype. **J Surg Oncol** 2009; 99:143–7.
- Shah MA, Ajani JA. Gastric cancer—an enigmatic and heterogeneous disease. **JAMA** 2010; 303:1753–4.
- Shekarian T, Valsesia-Wittmann S, Caux C, et al. Paradigm shift in oncology: targeting the immune system rather than cancer cells. **Mutagenesis** 2015; 30: 205–11.
- Shemirani AI, Haghighi MM, Zadeh SM, et al. Simplified MSI marker panel for diagnosis of colorectal cancer. **Asian Pac J Cancer Prev** 2011; 12:2101–4.

Smid M, Rodriguez-Gonzalez FG, Sieuwerts AM, et al. Breast cancer genome and transcriptome integration implicates specific mutational signatures with immune cell infiltration. **Nat Commun** 2016; 7:12910.

Stephens PJ, Tarpey PS, Davies H, et al. The landscape of cancer genes and mutational processes in breast cancer. **Nature** 2012; 486:400–4.

Stratton MR. Exploring the genomes of cancer cells: progress and promise. **Science** 2011; 331:1553–8.

Stratton MR, Campbell PJ, Futreal PA. The cancer genome. **Nature** 2009; 458:719–24.

Supek F, Lehner B. Differential DNA mismatch repair underlies mutation rate variation across the human genome. **Nature** 2015; 521:81–4.

Topalian SL, Hodi FS, Brahmer JR, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. **N Engl J Med** 2012; 366:2443–54.

Uemura N, Okamoto S, Yamamoto S, et al. Helicobacter pylori infection and the development of gastric cancer. **N Engl J Med** 2001; 345:784–9.

Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. **J Natl Cancer Inst** 2004; 96:261–8.

Wang S, Jia M, He Z, et al. APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. **Oncogene** 2018; 37:3924–36.

Wirtz HC, Muller W, Noguchi T, et al. Prognostic value and clinicopathological profile of microsatellite instability in gastric cancer. **Clin Cancer Res** 1998; 4:1749–54.

Wood RD, Mitchell M, Lindahl T. Human DNA repair genes, 2005. **Mutat Res** 2005; 577:275–83.

Yamamoto H, Adachi Y, Taniguchi H, et al. Interrelationship between microsatellite instability and microRNA in gastrointestinal cancer. **World J Gastroenterol** 2012; 18:2745–55.

Yamashita K, Sakuramoto S, Katada N, et al. Diffuse type advanced gastric cancer showing dismal prognosis is characterized by deeper invasion and emerging peritoneal cancer cell: the latest comparative study to intestinal advanced gastric cancer. **Hepatogastroenterology** 2009; 56:276–81.

Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. **Nat Commun** 2013; 4:2612.

Zhu L, Li Z, Wang Y, et al. Microsatellite instability and survival in gastric cancer: A systematic review and meta-analysis. **Mol Clin Oncol** 2015; 3:699–705.