

Um potencial biomarcador diagnóstico e prognóstico no câncer gástrico: uma análise *in silico* do gene GPNMB

Bianca de Andrade Lopes (biaalopes2003@gmail.com)
Universidade de Caxias do Sul (UCS)

Fernanda Pessi de Abreu (fpabreu1@ucs.br)
Universidade de Caxias do Sul (UCS)

Pedro Lenz Casa (plcasa@ucs.br)
Universidade de Caxias do Sul (UCS)

Marcos Vinícius Rossetto (rossettomarcos@gmail.com)
Universidade de Caxias do Sul (UCS)

Scheila de Avila e Silva (sasilva6@ucs.br)
Universidade de Caxias do Sul (UCS)

DOI: 10.18226/25253824.v8.n13.13

Submetido em: 30/05/2024 **Revisado em:** 10/09/2024 **Aceito em:** 10/09/2024

Resumo: O câncer gástrico é o quinto mais comum e a terceira maior causa de morte por câncer no mundo. No estágio inicial os pacientes podem ser assintomáticos ou não apresentar sintomas específicos, dificultando o diagnóstico. Um estudo prévio identificou 39 genes com potencialidade de utilização como biomarcadores no câncer gástrico, entre eles o gene GPNMB. Nesse contexto, este trabalho teve como objetivo explorar o GPNMB como biomarcador prognóstico e diagnóstico para o câncer gástrico. Foram utilizados dados de expressão extraídos dos repositórios Gene Expression Omnibus (GSE33335 e GSE54129) e do The Cancer Genome Atlas (TCGA-STAD). A aquisição dos dados, o processamento e aplicação das análises estatísticas foram realizados com uma ferramenta de desenvolvimento próprio. Os algoritmos K-means e árvore de decisão foram aplicados para determinar o potencial como biomarcador diagnóstico do gene, enquanto uma análise de sobrevida verificou a influência da expressão no prognóstico. A expressão de GPNMB foi maior em amostras de tecido tumoral quando comparadas com tecido não tumoral adjacente (NT). O K-means possibilitou a formação de grupos independentes com amostras normais e NT. De forma similar, com a árvore de decisão, as amostras foram corretamente classificadas em tecido normal e NT, de acordo com os valores de expressão. Além disso, as análises de sobrevida indicaram que a alta expressão do gene GPNMB está associada ao pior prognóstico. A pesquisa desenvolvida evidenciou o potencial do gene GPNMB como biomarcador para câncer gástrico, sendo esse demonstrado importante no desenvolvimento da doença.

Palavras-chaves: Câncer de Estômago, Expressão Diferencial, Inteligência Artificial, Gene Expression Omnibus, The Cancer Genome Atlas.

Abstract: Gastric cancer is the fourth most common and third deadliest cancer worldwide. Patients are usually asymptomatic or do not show specific symptoms during initial stages, which may hamper the diagnosis. A previous study identified 39 genes with biomarker potential in gastric cancer, among them the GPNMB gene. In this context, the objective of this study was to explore GPNMB as a prognostic and diagnostic biomarker for gastric cancer. Expression data was extracted from Gene Expression Omnibus (GSE33335 and GSE54129) and The Cancer Genome Atlas (TCGA-STAD). Data acquisition, preprocessing and statistical analyses were performed with an inhouse developed tool. The K-means and decision tree algorithms were applied for determining the potential of the gene as a diagnostic biomarker, whereas the survival analysis verified the influence of expression on prognosis. GPNMB expression was higher in tumoral tissue samples when compared to non-tumoral adjacent tissue (NT). K-means allowed formation of independent groups with normal and NT samples. Similarly, samples were correctly classified into normal and NT tissue groups with the decision tree according to expression values. Additionally, the survival analyses showed that the high expression of the GPNMB gene is associated with a worse prognosis. This research provided evidence on the potential of GPNMB as a biomarker for gastric cancer, given the gene demonstrated an important role in disease development.

Keywords: Stomach Cancer; Differential Expressio.; Artificial Intelligence, Gene Expression Omnibus, The Cancer Genome Atlas.

Introdução

A bioinformática, ou biocomputação, relaciona-se diretamente às descobertas da biologia molecular e auxilia no desenvolvimento do conhecimento científico da área. Dessa forma, as suas ferramentas possibilitam o estudo a nível molecular dos organismos vivos, análises a partir das quais também é possível identificar genes, proteínas e outras moléculas que auxiliem na elucidação de meios para diagnosticar anomalias e doenças. O destaque é dado à concepção de estratégias terapêuticas e o diagnóstico de tumores, como no caso das neoplasias [1].

A neoplasia maligna (câncer) é um grupo amplo e heterogêneo de patologias que iniciam com multiplicação desordenada e descontrolada de células. Essa proliferação fora de controle é uma característica das células cancerosas, promovendo a progressão do tumor e sua manutenção. Ao contrário das neoplasias benignas que tem crescimento mais lento e organizado [2,3]. As células podem tornar-se cancerosas por danos no DNA. Essas alterações no material genético podem transformar os proto-oncogenes, responsáveis pela regulação e o crescimento celular, em oncogenes causando descontrolo na proliferação celular. Além disso, as alterações no DNA podem ocasionar a desativação dos

genes supressores de tumor, que retardam a divisão celular ou levam a célula à morte no momento adequado [4].

Entre as neoplasias conhecidas está o câncer de estômago, ou câncer gástrico, é o quinto câncer mais comum e a terceira maior causa de morte por câncer no mundo [5]. Esta neoplasia está associada a fatores de risco como a infecção pela bactéria *Helicobacter pylori*, o baixo consumo de frutas e vegetais, o tabagismo e o etilismo. Além disso, o histórico familiar também é um fator que contribui significativamente para o desenvolvimento dessa doença [6].

O câncer gástrico não é uma doença única, mas um grupo heterogêneo de tumores com características morfológicas, histológicas e moleculares heterogêneas. A maioria dos cânceres gástricos são adenocarcinomas (95 % dos casos), que se iniciam no epitélio glandular [7,8]. Os adenocarcinomas gástricos podem ser classificados em: difuso, intestinal, e misto [8]. No estágio inicial, a doença pode ser assintomática ou não apresentar sintomas específicos, dificultando o diagnóstico precoce. Assim, os pacientes são diagnosticados em uma fase avançada da doença, reduzindo os índices de sucesso do tratamento e de sobrevivência dos pacientes [9,10]. Diante desse contexto, a utilização de abordagens mais eficazes que auxiliam na triagem inicial dos pacientes com câncer gástrico torna-se importante para possibilitar um melhor desfecho da doença [11].

As moléculas biológicas, como os biomarcadores, podem indicar se os processos em um organismo estão ocorrendo de forma normal ou anormal. Além disso, são utilizados em outros contextos clínicos, como para previsão do prognóstico, previsão da resposta do paciente a determinada terapia e monitoramento do estado da patologia [12].

Em um estudo prévio realizado por Rossetto (2021) foram identificados 39 genes diferencialmente expressos em câncer gástrico, o que indica uma potencialidade de utilização como biomarcadores [13]. Dentre esses, o gene *GNPMB* (glicoproteína B não metastática) é expresso na maioria das células e tecidos, e o incremento da expressão promove a proliferação, migração e a invasão de células tumorais [14]. Assim, devido às características do gene *GNPMB* e as sua potencialidade previamente relatada como biomarcador para o câncer gástrico, este estudo teve como objetivo utilizar técnicas de inteligência artificial (clusterização e árvore de decisão) conjuntamente com análise de sobrevivência a fim de explorar a utilização desse gene como biomarcador diagnóstico e prognóstico para neoplasia em questão.

Metodologia

Este trabalho foi desenvolvido com dados extraídos de repositórios públicos. Assim, não envolveu a coleta de novos dados em campo ou em laboratório.

Os dados de expressão gênica foram extraídos dos repositórios *Gene Expression Omnibus* (GEO) [15] e do *The Cancer Genome Atlas* (TCGA). Os conjuntos GSE33335 [16] e GSE54129 foram adquiridos no banco de dados GEO. O primeiro estudo contém, no total, 50 amostras, sendo 25 amostras não tumorais adjacentes (NT) e 25 amostras tumorais. Já o segundo estudo, possui no total 133 amostras, sendo 22 amostras NT e 111 amostras tumorais. No TCGA foram obtidos os dados do estudo TCGA-STAD [17], neste caso havia 415 amostras tumorais e 35 amostras de NT. Todos os dados são relacionados a pacientes diagnosticados com adenocarcinoma gástrico. A extração dos dados do GEO e TCGA, bem como o processamento e aplicação das análises estatísticas foram realizados por meio de uma ferramenta de desenvolvimento próprio que permite análise de expressão diferencial para identificação de potenciais biomarcadores [13].

Para análises dos dados provenientes do GEO foi utilizada a ferramenta *Orange v.3.26.0* [18]. Foram aplicadas duas técnicas de mineração de dados, sendo os algoritmos *K-means* e a *Árvore de Decisão*. O *K-means* possibilita o agrupamento das amostras de cada estudo segundo similaridade na expressão do gene *GNPMB*. O critério levado em consideração na análise dos resultados foi a pureza dos clusters de acordo com o tipo de amostra prevalente (tumoral ou NT). Já o algoritmo de *Árvore de Decisão* permite a classificação das amostras em tumoral e NT utilizando como critério os valores de expressão do gene *GNPMB*. As árvores de decisão geradas foram avaliadas através da validação cruzada segundo as métricas: (i) AUC (área embaixo da curva); (ii) exatidão (proporção de exemplos classificados corretamente); (iii) precisão (proporção de verdadeiros positivos classificados corretamente); (iv) revocação (proporção de verdadeiros positivos entre todos os exemplos positivos); e (v) F1 (média harmônica ponderada entre medidas de precisão e revocação).

Os dados provenientes do TCGA foram submetidos à análise de sobrevivência de acordo com a categoria histopatológica (intestinal e difuso). Essa análise foi também realizada através da ferramenta de desenvolvimento próprio. Os valores *p* menores que 0,05 foram tomados como estatisticamente significativos.

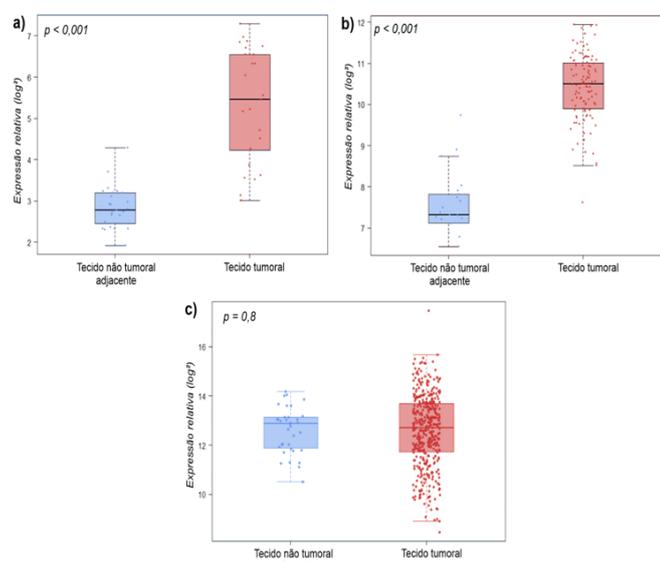
Resultados

Como resultado da análise de expressão diferencial, o gene *GNPMB* apresentou diferença estatisticamente significativa entre amostras tumorais e NT em dois dos três conjuntos de dados (figura 1). Para os conjuntos GSE33335 e GSE54129 o gene apresentou uma expressão mais alta em tecido tumoral ($p < 0,001$), enquanto que os dados provenientes do estudo TCGA-STAD não evidenciaram diferença de expressão entre os dois tipos de tecido ($p = 0,8$).

Os dados do GEO (GSE33335 e GSE54129) foram submetidos a simulações com valores de *K* variando de 2 a 5. Os resultados da análise de clusterização estão apresentados nas

tabelas 1 e 2. Para o GSE33335 o valor de K igual a 3 proporcionou melhor resultado (pureza média de 90,92 %). Já para o GSE54129 o melhor resultado (pureza média de 95,56 %) foi atingido com valor de K igual a 4. Assim, é perceptível que os valores de expressão do gene *GNMB* possibilitaram o agrupamento das amostras de acordo com o tecido (tecido normal e tecido tumoral).

Figura 1: Comparação da expressão gênica do gene *GNMB* entre amostras tumorais e tecido não tumoral adjacente. a) Conjunto de dados do estudo GSE33335; b) Conjunto de dados do estudo GSE54129; c) Conjunto de dados do estudo TCGA-STAD.



Fonte: autores.

Tabela 1: Resultados da análise de clusterização K -means com valores de expressão do gene *GNMB* do conjunto de dados GSE33335. Os resultados para o melhor valor de K estão realçados em negrito.

Valor de K	Cluster	Normal	Tumor	Pureza dos clusters	Pureza média
2	C1	0	16	100,00 %	86,76 %
	C2	25	9	73,53 %	
3	C1	0	11	100,00 %	90,92 %
	C2	24	5	82,76 %	
	C3	1	9	90,00 %	
4	C1	0	11	100,00 %	87,85 %
	C2	10	6	62,50 %	
	C3	1	8	88,89 %	
	C4	14	0	100,00 %	
5	C1	0	5	100,00 %	90,00 %
	C2	10	5	66,67 %	
	C3	1	5	83,33 %	
	C4	0	10	100,00 %	
	C5	14	0	100,00 %	

Fonte: autores.

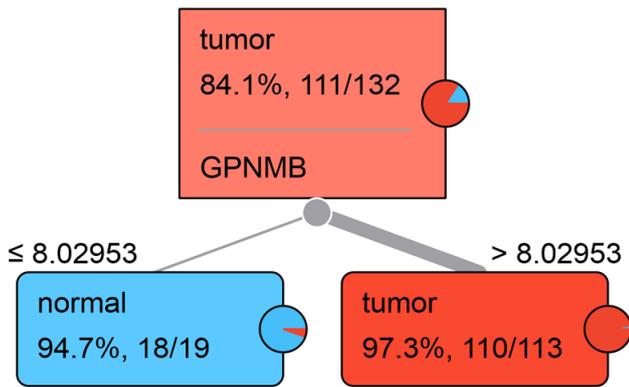
Também foram realizadas análises com algoritmo de “Árvore de Decisão”. Na figura 2, foi apresentada a Árvore de Decisão gerada para o estudo GSE54129, sendo que as amostras com valores de expressão menor ou igual a 8,02 foram classificadas como normais, por outro lado, amostras com valores maiores que 8,02 foram classificadas como tumorais. Na figura 3, está sendo apresentada a Árvore de Decisão gerada para o estudo GSE33335, sendo que as amostras com valores de expressão menor ou igual a 3,30 foram classificadas normais, enquanto amostras com valores maiores que 3,30 foram classificadas como tumorais. As métricas utilizadas para avaliação das duas árvores tiveram elevado desempenho (tabela 3), ratificando o potencial dos valores de expressão do gene *GNMB* para classificação das amostras entre normais e tumorais.

Tabela 2: Resultados da análise de clusterização K -means com valores de expressão do gene *GNMB* do conjunto de dados GSE54129. Os resultados para o melhor valor de K estão realçados em negrito.

Valor de K	Cluster	Normal	Tumor	Pureza dos clusters	Pureza média
2	C1	1	103	99,04 %	85,23 %
	C2	20	8	71,43 %	
3	C1	3	46	93,88 %	93,20 %
	C2	18	3	85,71 %	
	C3	0	62	100,00 %	
4	C1	3	21	87,50 %	95,56 %
	C2	0	40	100,00 %	
	C3	18	1	94,74 %	
	C4	0	49	100,00 %	
5	C1	18	1	94,74 %	94,99 %
	C2	0	39	100,00 %	
	C3	1	31	96,88 %	
	C4	2	10	83,33 %	
	C5	0	30	100,00 %	

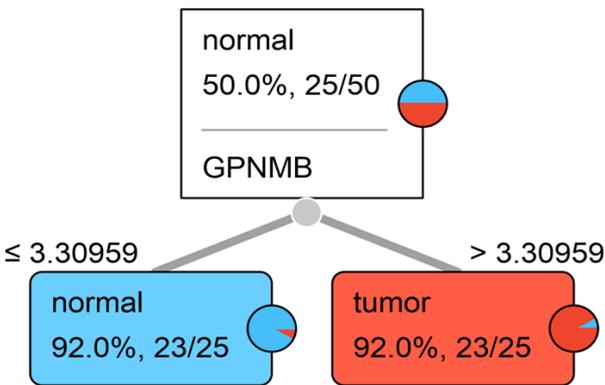
Fonte: autores.

Figura 3: Resultado da análise de Árvore de Decisão com valores de expressão do gene *GPNMB* do conjunto de dados GSE54129.



Fonte: autores.

Figura 4: Resultado da análise de Árvore de Decisão com valores de expressão do gene *GPNMB* do conjunto de dados GSE33335.



Fonte: autores.

Tabela 3: Métricas utilizadas para validação cruzada das árvores de decisão geradas. Abreviações: AUC, área embaixo da curva; CA, exatidão.

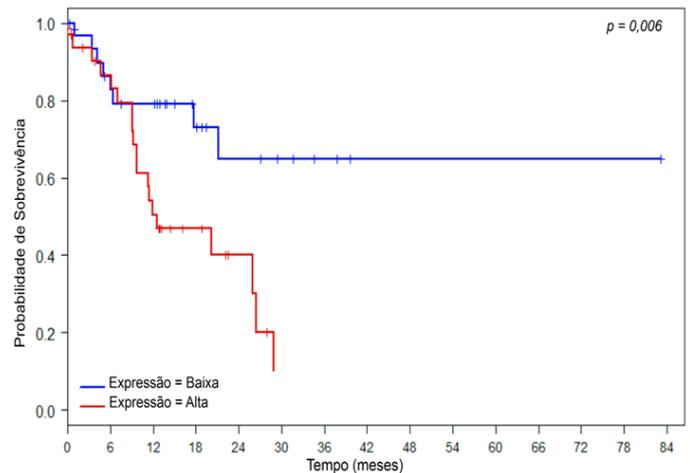
Estudos	AUC	CA	Precisão	Revocação	F1
GSE54129	0.893	0.955	0.953	0.955	0.954
GSE54129	0.862	0.900	0.901	0.900	0.900

Fonte: autores.

Em relação às análises de sobrevida com os dados do TCGA, os resultados apontam para uma diferença na probabilidade de sobrevivência dos grupos de pacientes dependendo da classificação histopatológica da neoplasia (tipo difuso ou intestinal). Para o câncer gástrico do tipo difuso (figura 5), houve uma diferença estatisticamente significativa entre a sobrevida dos pacientes que apresentaram alta e baixa expressão. Nesse caso, os pacientes com baixa expressão do gene *GPNMB* tiveram uma maior taxa de sobrevida ($p = 0,006$). No entanto, os níveis de expressão de *GPNMB* não foram bons preditores para a sobrevida dos pacientes com câncer gástrico do tipo intestinal ($p = 0,683$), visto que as

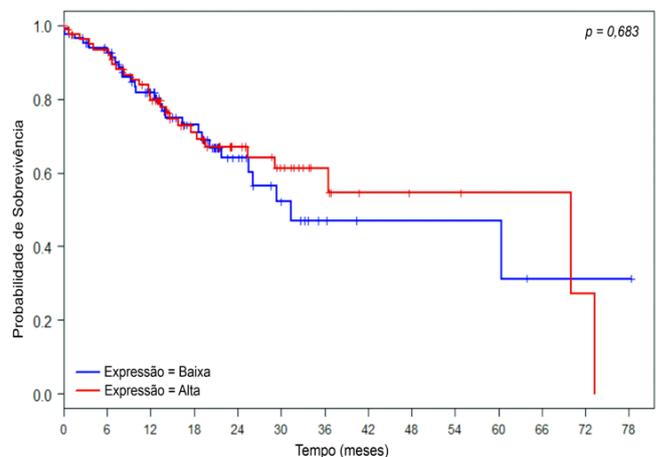
curvas de sobrevida dos pacientes com alta e baixa expressão não tiveram diferença estatisticamente significativa (figura 6).

Figura 5: Comparação da sobrevida baseada na expressão de *GPNMB* para o câncer gástrico do tipo difuso.



Fonte: autores.

Figura 6: Comparação da sobrevida baseada na expressão de *GPNMB* para o câncer gástrico do tipo intestinal.



Fonte: autores.

Discussão

O gene *GPNMB* codifica uma glicoproteína super expressa em uma variedade de cânceres. Esse gene aparenta ter papel imunossupressor, o que contribui para o crescimento tumoral e metástase [19]. No câncer de cabeça e pescoço, *GPNMB* foi altamente expresso quando comparado com displasia e tecido de mucosa normal [20].

De forma similar aos estudos com outros tipos de câncer, a análise desse estudo revelou o gene *GPNMB* como um potencial biomarcador diagnóstico para o câncer gástrico. Ao comparar os valores de expressão do gene *GPNMB* entre amostras de

tecido não tumoral adjacente e tecido tumoral, os dados do GEO apresentaram diferença estatisticamente significativa. Por outro lado, os valores de expressão entre os dois tipos de tecido não foram estatisticamente significativos para os dados do TCGA, podendo ser explicado pela discrepância do número de amostras de cada grupo. Ao considerar os valores de expressão do gene *GPNMB* foi possível a formação de grupos independentes com amostras normais e tumorais na análise de clusterização. Além disso, os dados de expressão também permitiram a classificação das amostras, sendo valores mais altos de expressão encontrados em tecido tumoral e os valores mais baixos de expressão em tecidos não tumorais adjacentes, conforme evidenciado pelas árvores de decisão modeladas.

Nesse estudo, na análise de sobrevida, a alta expressão do gene *GPNMB* foi correlacionado com pior prognóstico no câncer gástrico do tipo difuso. Ren e colaboradores (2020) também investigaram o papel do gene *GPNMB* no câncer gástrico e relacionaram altos níveis de expressão com a predição de um pior prognóstico. A proteína codificada por *GPNMB* tem a capacidade de recrutar as células imunossupressoras e promover a exaustão das células imunes através do eixo de sinalização PI3K/AKT/CCL4, atuando como um mediador na progressão do câncer gástrico [21].

Adicionalmente, a alta expressão de *GPNMB* também é correlacionada com mau prognóstico e menores taxas de sobrevida em pacientes com outros tipos de câncer [19]. No câncer de cabeça e pescoço, a um alto nível de expressão de *GPNMB*, sendo associado a um pior prognóstico [20]. No câncer de mama triplo negativo, a expressão de *GPNMB* é prevalente e pode ser utilizado no prognóstico dos pacientes, visto que a superexpressão foi associada a um pior prognóstico e metástase [22]. Além do seu potencial como biomarcador diagnóstico e prognóstico, o gene *GPNMB* também foi estudado como alvo terapêutico para o câncer de mama triplo negativo [23].

Considerações finais

Os algoritmos utilizados possibilitaram a identificação do gene *GPNMB* como potencial biomarcador diagnóstico para o câncer gástrico, visto que a expressão diferencial entre os tecidos tumorais e não tumorais adjacentes possibilitou o agrupamento e a classificação das amostras. As amostras de tecido tumoral apresentaram valores elevados de expressão, enquanto as amostras não tumorais adjacentes apresentaram valores mais baixos de expressão, indicando que esse gene possui um papel no desenvolvimento do câncer gástrico. Em geral, os pacientes com essa neoplasia são diagnosticados tardiamente e a identificação de potenciais biomarcadores diagnósticos pode auxiliar no diagnóstico precoce. Além disso, as análises de sobrevida indicaram que pacientes com alta expressão do gene *GPNMB* possuem um pior prognóstico do câncer gástrico do tipo difuso.

Por fim, este trabalho pode motivar estudos futuros, relacionados a validação clínica do gene *GPNMB* como biomarcador. Além disso, pesquisas adicionais são necessárias para explorar as vias moleculares através das quais o *GPNMB* influencia o desenvolvimento e a progressão do câncer gástrico. Apesquisa sobre a possibilidade de inibir o *GPNMB* como uma estratégia terapêutica é um campo promissor na área oncológica, que pode abrir novas opções de tratamento para pacientes com câncer gástrico.

Referências

- [1] de ARAÚJO, N. D., de FARIAS, R. P., PEREIRA, P. B., de FIGUEIRÊDO, F. M., de MORAIS, A. M. B., SALDANHA, L. C., & GABRIEL, J. E. (2008). A era da bioinformática: seu potencial e suas implicações para as ciências da saúde. *Estudos de biologia*, 30(70/72): 143-8 DOI: <https://doi.org/10.7213/rev.v30i70/72.22819>.
- [2] INSTITUTO Nacional de Câncer (INCA). O que é câncer? Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/o-que-e-cancer>. Acesso em: 21 set. 2022.
- [3] VOGELSTEIN, B., & KINZLER, K. W. (2015). The path to cancer—three strikes and you’re out. *N Engl J Med*, 373 (20), 1895-1898. DOI: <https://www.nejm.org/doi/10.1056/NEJMp1508811>
- [4] Ward, L. S. (2002). Entendendo o processo molecular da tumorigênese. *Arquivos Brasileiros de Endocrinologia & Metabologia*, 46, 351-360. DOI: <https://doi.org/10.1590/S0004-27302002000400006>
- [5] BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R. L., TORRE, L. A., & JEMAL, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68 (6), 394-424. DOI: <https://doi.org/10.3322/caac.21492>
- [6] MACHLOWSKA, J., BAJ, J., SITARZ, M., MACIEJEWSKI, R., & SITARZ, R. (2020). Gastric cancer: epidemiology, risk factors, classification, genomic characteristics and treatment strategies. *International journal of molecular sciences*, 21(11), 4012. DOI: <https://doi.org/10.3390/ijms21114012>
- [7] Yakirevich, E., & Resnick, M. B. (2013). Pathology of gastric cancer and its precursor lesions. *Gastroenterology Clinics*, 42 (2), 261-284. DOI: <https://doi.org/10.1016/j.gtc.2013.01.004>
- [8] VAN CUTSEM, E., SAGAERT, X., TOPAL, B., HAUSTERMANS, K., & PRENEN, H. (2016). Gastric

cancer. *The Lancet*, 388 (10060), 2654-2664. DOI: [http://dx.doi.org/10.1016/S0140-6736\(16\)30354-3](http://dx.doi.org/10.1016/S0140-6736(16)30354-3)

[9] PASECHNIKOV, V., CHUKOV, S., FEDOROV, E., KIKUSTE, I., & LEJA, M. (2014). Gastric cancer: prevention, screening and early diagnosis. *World journal of gastroenterology*: WJG, 20 (38), 13842. <http://dx.doi.org/10.3748/wjg.v20.i38.13842>

[10] LEE, J. Y., KIM, H. I., KIM, Y. N., HONG, J. H., ALSHOMIMI, S., AN, J. Y., & KIM, C. B. (2016). Clinical significance of the prognostic nutritional index for predicting short-and long-term surgical outcomes after gastrectomy: a retrospective analysis of 7781 gastric cancer patients. *Medicine*, 95 (18), e3539. DOI: <https://doi.org/10.1097/md.0000000000003539>

[11] MATSUOKA, T., & YASHIRO, M. (2018). Biomarkers of gastric cancer: Current topics and future perspective. *World journal of gastroenterology*, 24 (26), 2818. DOI: <https://doi.org/10.3748%2Fwjg.v24.i26.2818>

[12] HENRY, N. L., & HAYES, D. F. (2012). Cancer biomarkers. *Molecular oncology*, 6 (2), 140-146. DOI: <https://doi.org/10.1016/j.molonc.2012.01.010>

[13] ROSSETTO, M. V. (2019). *Análise de expressão diferencial de genes: uma solução computacional para identificação de biomarcadores de tumores gástricos em humanos*. Dissertação, Universidade de Caxias do Sul <https://repositorio.ucs.br/11338/8780>

[14] TAYA, M., & HAMMES, S. R. (2018). Glycoprotein non-metastatic melanoma protein B (GPNMB) and cancer: a novel potential therapeutic target. *Steroids*, 133, 102-107. DOI: <https://doi.org/10.1016/j.steroids.2017.10.013>

[15] EDGAR, R., DOMRACHEV, M., & LASH, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids research*, 30 (1), 207-210. DOI: <https://doi.org/10.1093/nar/30.1.207>

[16] CHENG, L., YANG, S., YANG, Y., ZHANG, W., XIAO, H., GAO, H. & ZHANG, Q. (2012). Global gene expression and functional network analysis of gastric cancer identify extended pathway maps and GPRC5A as a potential biomarker. *Cancer letters*, 326(1), 105-113. DOI: <https://doi.org/10.1016/j.canlet.2012.07.031>

[17] TCGA. Cancer Genome Atlas Research Network. (2014). Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 513, 202-209.

[18] DEMŠAR, J., CURK, T., ERJAVEC, A., GORUP, Č., HOČEVAR, T., MILUTINOVIČ, M., & ZUPAN, B. (2013).

Orange: data mining toolbox in Python. *The Journal of machine Learning research*, 14 (1), 2349-2353. Disponível em: <https://www.jmlr.org/papers/volume14/demsar13a/demsar13a.pdf>

[19] LAZARATOS, A. M., ANNIS, M. G., & SIEGEL, P. M. (2022). GPNMB: a potent inducer of immunosuppression in cancer. *Oncogene*, 41 (41), 4573-4590. DOI: <https://doi.org/10.1038/s41388-022-02443-2>

[20] LI, H., XIAO, Y., WU, C. C., YANG, L. L., CAO, L. Y., CHEN, D. R. & SUN, Z. J. (2019). High expression of GPNMB predicts poor prognosis in head and neck squamous cell carcinoma. *Histology and Histopathology*, 34(7). DOI: <https://doi.org/10.14670/hh-18-084>

[21] REN, F., ZHAO, Q., LIU, B., SUN, X., TANG, Y., HUANG, H. & MI, Y. (2020). Transcriptome analysis reveals GPNMB as a potential therapeutic target for gastric cancer. *Journal of Cellular Physiology*, 235(3), 2738-2752. DOI: <https://doi.org/10.1002/jcp.29177>

[22] HUANG, Y. H., CHU, P. Y., CHEN, J. L., HUANG, C. T., HUANG, C. C., TSAI, Y. F., & LIU, C. Y. (2021). Expression pattern and prognostic impact of glycoprotein non-metastatic B (GPNMB) in triple-negative breast cancer. *Scientific reports*, 11 (1), 12171. DOI: <https://doi.org/10.1038/s41598-021-91588-3>

[23] MARIC, G., ROSE, A. A., ANNIS, M. G., & SIEGEL, P. M. (2013). Glycoprotein non-metastatic b (GPNMB): A metastatic mediator and emerging therapeutic target in cancer. *OncoTargets and therapy*, 9 (6) 839-852. DOI: <https://doi.org/10.2147/OTT.S44906>