



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

ANÁLISE DA EXPRESSÃO GÊNICA DE PACIENTES COM ADENOCARCINOMA
GÁSTRICO ASSOCIADO AO VÍRUS EPSTEIN-BARR

VALÉRIA CRISTIANE SANTOS DA SILVA

Belém

2024

VALÉRIA CRISTIANE SANTOS DA SILVA

**ANÁLISE DA EXPRESSÃO GÊNICA DE PACIENTES COM
ADENOCARCINOMA GÁSTRICO ASSOCIADO AO VÍRUS EPSTEIN-
BARR**

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia, da Universidade Federal do Pará, como requisito parcial para obtenção do título de mestre em Biotecnologia.

Orientador: Prof. Dr. Fabiano Cordeiro Moreira

Belém

2024

VALÉRIA CRISTIANE SANTOS DA SILVA

ANÁLISE DA EXPRESSÃO GÊNICA DE PACIENTES COM
ADENOCARCINOMA GÁSTRICO ASSOCIADO AO VÍRUS EPSTEIN-BARR

Dissertação apresentada ao Programa de Pós-Graduação em
Biotecnologia, da Universidade Federal do Pará, como requisito
parcial para obtenção do título de mestre em Biotecnologia.

Orientador: Prof. Dr. Fabiano Cordeiro Moreira

Belém - PA, 12 de agosto de 2024

Conceito: _____

Banca Examinadora:

Prof. Dr. Fabiano Cordeiro Moreira (Orientador)
Núcleo de Pesquisas em Oncologia / UFPA

Prof. Dr. Samir Mansour Casseb
Núcleo de Pesquisas em Oncologia/UFPA

Prof. Dr. Rafael de Azevedo Baraúna
Instituto de Ciências Biológicas/ UFPA

Prof. Dr. André Salim Khayat
Instituto de Ciências Biológicas/UFPA

Belém

2024

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo
com ISBD Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

- S586a Silva, Valeria Cristiane Santos da.
Análise da expressão gênica de pacientes com adenocarcinoma gástrico
associado ao vírus Epstein-Barr / Valeria Cristiane Santos da Silva, . —
2024.
63 f. : il. color.
- Orientador(a): Prof. Dr. Fabiano Cordeiro Moreira Dissertação
(Mestrado) - Universidade Federal do Pará,
Instituto de Ciências Biológicas, Programa de Pós-Graduação em
Biotecnologia, Belém, 2024.
1. vírus Epstein-Barr. 2. câncer gástrico. 3. RNA-seq. 4. HHV-
4. I. Título.

CDD 610

INSTITUIÇÕES PARTICIPANTES E FINANCIADORAS

INSTITUIÇÕES PARTICIPANTES:

1. Laboratório de Genética Humana e Médica (LGHM) – Instituto de Ciências Biológicas/Universidade Federal do Pará (UFPA)
2. Núcleo de Pesquisas em Oncologia/Universidade Federal do Pará (UFPA)
3. Hospital Universitário João de Barros Barreto (HUJBB)/Universidade Federal do Pará (UFPA)
4. Hospital Ophir Loyola (HOL)
5. Laboratório de Anatomia Patológica do Complexo Hospitalar Universitário do Hospital Universitário João de Barros Barreto/ Universidade Federal do Pará
6. Instituto Senai de Inovação em Tecnologias Minerais
7. Instituto Tecnológico Vale

INSTITUIÇÕES FINANCIADORAS:

1. Coordenação de Aperfeiçoamento Pessoal de Ensino Superior (CAPES)
2. Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA)
3. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

Aos meus pais, pela oportunidade de realizar mais um sonho.

AGRADECIMENTOS

À Deus, por ter me guiado até aqui com graça e proteção.

Aos meus pais José Afonso e Francisca Silva, que são fonte de inspiração e incentivo, por sempre acreditarem em mim e me apoiarem incondicionalmente. Aos meus irmãos Rogério, Felipe e Edson pelo amor, cuidado e apoio ao longo da vida. Nenhuma palavra vai ser o bastante para externalizar toda a minha gratidão e amor a vocês.

À minha avó Maria Antônia, por todas as palavras de incentivo, amor e por me inspirar cada dia mais a ser uma mulher forte e sonhadora. A minha prima e melhor amiga, Lady Anne, por sempre acreditar em mim e me incentivar a realizar meus sonhos, mesmo quando tudo parecia impossível.

Ao meu orientador Prof. Drº Fabiano Moreira, por ser fonte de inspiração, por todos os ensinamentos e incentivos ao longo desses anos.

À equipe de Bioinformática do NPO, por toda a ajuda, palavras de incentivos, aprendizado e *happy hours* compartilhados. O caminho acadêmico tem sido mais leve ao lado de vocês.

A meus amigos da UFPA, Beatriz, Davi, Giovane, Gustavo, Joaquim, Naiana e Thiago, por se tornarem um ponto de apoio essencial na minha vida, por todo aprendizado e por sempre estarem presentes.

Aos meus amigos, que sempre me acompanham em todas as etapas da vida, ouviram todos os meus lamentos e vibraram com minhas conquistas. Ana Paula, Bianca, Diego, Emily, Fernanda, Ian, Márcia, Rayssa, Samuel e W. Junior, obrigada por tudo.

Às instituições e autoridades que concederam autorizações para realização dessa pesquisa.

À Capes pelo apoio financeiro na concessão da bolsa de estudo.

RESUMO

A infecção pelo vírus Epstein-Barr (EBV) é um dos fatores de risco para o câncer gástrico (CG). O Epstein-Barr é um vírus que possui atividade oncogênica e foi o primeiro a ser associado a doenças malignas como linfomas e carcinomas. Normalmente, a detecção de EBV pode ser feita por *hibridização in situ* e atualmente por meio de utilização de técnicas de sequenciamento de RNA para identificar a presença de genes virais em amostras. Nesse sentido, buscando compreender melhor sobre o subtipo positivo para EBV, este trabalho propôs uma caracterização molecular por meio de sequenciamento de RNA a partir de 76 amostras tumorais de pacientes diagnosticados com CG. Para isso, foi realizado o sequenciamento de RNA total das amostras e para a caracterização molecular foi utilizado o *software Kraken2*. Das 76 amostras, 8 foram consideradas positivas segundo o método adotado. Em seguida, para compreender os diferentes mecanismos pelo qual o EBV pode estar atuando no CG, foram analisados os padrões de expressão gênica humana das amostras classificadas como positivas e negativas. Em nosso estudo, há cerca de 834 genes diferencialmente expressos, dos quais 92 apresentaram $AUC > 0.85$. Esses genes estão associados a progressão tumoral, metabolismo celular e imunidade inata e adaptativa. Além disso, foram avaliados os genes virais expressos nas amostras consideradas positivas, e encontramos manifestação tanto genes da fase lítica quanto da fase latente. Por fim, nosso estudo apresenta uma estratégia eficiente para a classificação molecular do CG subtipo EBV positivo baseada em NGS e mostra os efeitos do EBV na expressão gênica humana.

Palavras-chave: vírus Epstein-Barr, cancer gástrico, RNA-seq, HHV-4

ABSTRACT

Infection with the Epstein-Barr virus (EBV) is one of the risk factors for gastric cancer (GC). Epstein-Barr is a virus with oncogenic activity and was the first to be associated with malignant diseases such as lymphomas and carcinomas. Thus, EBV detection can be performed both by in situ hybridization and, currently, by using RNA sequencing techniques to identify the presence of viral genes in samples. In this context, aiming to better understand the EBV-positive subtype, this study proposed a molecular classification based on RNA sequencing of 76 tumor samples from patients diagnosed with GC. RNA sequencing of the samples was performed, and molecular classification was done using the Kraken2 software. Of the 76 samples, 8 were considered positive according to the adopted method. Subsequently, to understand the different mechanisms by which EBV might be acting in gastric cancer, we analyzed the patterns of human gene expression in samples classified as positive and negative. In our study, there are approximately 834 differentially expressed genes, of which 92 have an AUC greater than 0.85. These genes are associated with tumor progression, cellular metabolism, and innate and adaptive immunity. Additionally, viral genes expressed in the positive samples were evaluated, and we found manifestations of both lytic phase and latent phase genes. Finally, our study presents an efficient strategy for the molecular classification of EBV-positive gastric cancer based on NGS and shows the effects of EBV on human gene expression.

Keywords: Epstein-Barr virus, gastric adenocarcinoma, RNA-seq, HHV-4.

LISTA DE ILUSTRAÇÕES

Figura 1 - <i>Hallmarks of Cancer</i> com foco nos polimorfismos microbianos.....	17
Figura 2 - Estrutura do vírus Epstein-Barr	19
Figura 3 - Ciclo de replicação do EBV.	21
Figura 4 - <i>Volcano plot</i> apresentando os genes DE entre amostras positivas e negativas para o EBV.	31
Figura 5 - Gráficos de curva ROC representando a AUC por gene.....	32
Figura 6 - <i>Heatmap</i> dos padrões de expressão dos genes DE	34
Figura 7 - Ontologia Gênica para processos biológicos.....	35
Figura 8 - Expressão dos genes virais de EBV.....	36
Figura 9 - Expressão dos genes virais do EBV por amostras.....	37
Figura 10 - Correlação entre os genes DE com AUC > 0,85 e os genes virais expressos nas amostras positivas.....	38

LISTA DE TABELAS E QUADROS

Quadro 1 - Expressão gênica do ciclo de latência do EBV.	22
Quadro 2 - Comparação das amostras positivas entre método de RNA-Seq e IHC	30
Tabela 1 - Genes humanos DE com AUC > 0.85.	33

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Hipótese	13
2 OBJETIVOS	15
2.1 Objetivo Geral	15
2.2 Objetivos Específicos	15
3 REVISÃO LITERÁRIA	16
3.1 Câncer Gástrico	16
3.2 Human gammaherpesvirus 4 (Epstein-Barr)	19
3.2.1 Replicação viral	20
3.2.1 Alterações genéticas em EBVaCG	23
3.3 Tratamentos e Diagnósticos	24
3.4 Metatranscriptoma por RNA-Seq	26
4 METODOLOGIA	27
4.1 Aspectos éticos e obtenção das amostras	27
4.2 Extração de RNA total	27
4.3 Construção das bibliotecas e sequenciamento de RNA	27
4.4 Controle de qualidade das amostras	28
4.5 Classificação Molecular de EBVaCG	28
4.7 Alinhamento das amostras com o genoma humano	28
4.7 Análises Multivariadas	28
4.7.1. Análise Diferencial dos Genes Humanos	28
4.7.2. Análises da Característica de Operação do Receptor	29
4.7.3. Análises de enriquecimento funcional dos genes	29
4.8 Análise de Expressão de Genes Virais	29
4.8.1. Correlação da expressão dos genes humanos DE com genes virais	29
5 RESULTADOS	30
5.1 Classificação Molecular do EBV	30
5.2 Análise da Expressão Gênica Diferencial	31
5.2.1. Análise de enriquecimento funcional dos genes DE	34
5.3 Análise da Expressão Viral	35
5.3.1. Correlação da expressão dos genes humanos DE com genes virais	37
6 DISCUSSÃO	39
7 CONCLUSÃO	50
REFERÊNCIAS BIBLIOGRÁFICAS	51
ANEXO I	59
ANEXO II	62

1 INTRODUÇÃO

O câncer gástrico (CG) ocupa o quinto lugar em números de casos e quarto em número de mortes por câncer no cenário mundial (Sung et al, 2021). No Brasil, o tipo de tumor gástrico mais recorrente é o adenocarcinoma gástrico (AG), responsável por 95% das incidências, com enfoque nas regiões norte e nordeste do país, onde há maior índice de casos (Inca, 2020). Conforme Curado e colaboradores (2019), a maior incidência de CG foi observada em Belém, onde as taxas foram três vezes maiores que as relatadas nas cidades de São Paulo e Fortaleza.

As principais causas para o surgimento e desenvolvimento do CG, estão associadas a falhas nos mecanismos reguladores do genoma (Grady *et al.*,2020), fatores ambientais, tais como hábitos alimentares e consumo exagerado de álcool e tabaco e a infecções causadas por microrganismos, tais como *Helicobacter Pylori* (*H.pylori*) e vírus Epstein-Barr (EBV) (Fang, 2015).

Para melhorar a compreensão acerca do CG, há classificações feitas de acordo com a histologia e as características moleculares. A classificação baseada nas características moleculares foi proposta pelo *The Cancer Genome Atlas* (TCGA) no ano de 2014 e está dividida em 4 subtipos, são eles: tumores com instabilidade de microssatélites (MSI), estáveis genomicamente (GE), tumores com instabilidade cromossômica (CIN) e por fim aqueles positivos para EBV ou EBV associado ao CG (EBVaCG) (Sunakawa; Lenz, 2015).

Os tumores associados ao EBV ocorrem devido à infecção pelo vírus Epstein-Barr, que desencadeia alterações epigenéticas como a hipermetilação acentuada do DNA, mutações nos genes *PIK3CA*, *BCOR* e *ARID1A*, além da amplificação de *JAK2* e superexpressão de *PD-L1* e *PD-L2*. Há também o silenciamento do gene *CDKN2A*, responsável por codificar duas proteínas que atuam como supressores tumorais. No estudo realizado pelo TCGA, a taxa de incidência do subtipo EBVaCG em tumores gástricos foi de aproximadamente 9% (Sunakawa; Lenz, 2015).

O Epstein-Barr (*Human gammaherpesvirus 4*) é um vírus da família *Herpesviridae*, sendo o primeiro vírus a ser associado a doenças malignas devido a sua atividade oncogênica. Foi associado a neoplasias como: câncer nasofaríngeo, linfoma de Hodgkin, linfoma de Burkitt e também o câncer gástrico (Shinozaki-Ushiku *et al.*, 2015). À luz dessa informação, vale trazer à tona que a infecção por EBV induz modificações epigenéticas que desempenham um papel importante na oncogênese e na progressão de tumores, por isso torna-se fundamental analisar

as expressões gênicas do vírus quando associado ao adenocarcinoma gástrico para identificar suas particularidades (Borozan *et al.*, 2018).

1.1 Hipótese

Para este estudo, o foco buscou entender a respeito dos AG classificados como EBVaCG, visto que identificar a presença de EBV em amostras de AG e analisar suas influências no processo de carcinogênese apresentam importante relevância clínica, pois além das particularidades moleculares e histológicas, pacientes com diagnóstico positivo para esta infecção viral demandam diferentes condutas terapêuticas. Nesse sentido, é válido destacar a importância de estudos que caracterizem os microbiomas virais para a compreensão da biologia do AG (Barra, 2021).

A presença do EBV foi associada pela primeira vez a um paciente de CG por meio da utilização de uma análise de PCR em 1990. Entretanto, apenas quando técnicas de *hibridização in situ* foram utilizadas para a detecção de pequenos RNAs codificados por EBV (EBER's), foi possível facilitar a identificação do vírus em amostras do adenocarcinoma gástrico (Iizasa *et al.*, 2012).

Ainda que já existam métodos de detecção para o EBV, não é possível analisar, por meio de testes imuno-histoquímicos, as alterações causadas pelo vírus no hospedeiro, além de não possibilitar a caracterização de outras similaridades moleculares que podem ocorrer simultaneamente no tumor (Bai *et al.*, 2022). Dessa maneira, busca-se uma classificação mais abrangente a respeito do vírus e suas influências na carcinogênese por meio de análise de sequenciamento de RNA, uma vez que o sequenciamento de RNA total consiste em detectar transcritos, bem como múltiplas formas de RNA não codificantes, para identificar e quantificar a expressão gênica destes.

Abordagens computacionais recentes possibilitam a análise taxonômica e funcional a partir de dados de sequenciamento de RNA em amostras de câncer, resultando na identificação de possíveis microrganismos associados à doença (Liu *et al.*, 2019). Dessa forma, utilizar métodos de RNA-seq para identificar os subtipos de AG de acordo com a classificação molecular se torna fundamental para a compreensão da heterogeneidade do câncer e das complexidades do EBV (Shinozaki-Ushiku *et al.*, 2015).

O ciclo de infecção do EBV é dividido em ciclo lítico e ciclo latente. As expressões virais do Epstein-Barr mais incidentes no AG são de proteínas que pertencem ao ciclo latente

como: *EBER's*, *EBNA1*, *BART*, *BARF1* e em alguns casos há a expressão de *LMP1* e *LMP2A*. Entretanto, em amostras do TCGA, foram encontrados genes que são manifestadas no ciclo lítico e estão relacionados à multiplicação e à manutenção de novas cópias do vírus (Borozan *et al.*, 2018). Portanto, a investigação da expressão gênica do vírus nas células tumorais se faz necessária para avaliar que parte do ciclo do vírus é predominante nas amostras tumorais, como isso afeta o desenvolvimento do câncer e qual impacto teria para o prognóstico do paciente.

Além disso, com os resultados provenientes do método de classificação, este trabalho objetiva avaliar como a regulação da expressão gênica humana é afetada na presença do vírus. Também, faz-se necessário comparar com o que já há descrito na literatura, para investigar se há diferentes influências do EBV em pacientes do estado do Pará, visto que Naseem e colaboradores (2018) sustentam a ideia de que polimorfismos genéticos podem causar diferentes cepas de EBV que podem contribuir para a variância geográfica da taxa de incidência do subtipo EBVaCG.

2 OBJETIVOS

2.1 Objetivo Geral

- Identificar a presença do vírus Epstein-Barr e investigar o seu efeito na expressão gênica humana em amostras de câncer gástrico a partir do sequenciamento de RNA.

2.2 Objetivos Específicos

- Desenvolver um método *in-silico* para classificar molecularmente o vírus Epstein-Barr nas amostras de adenocarcinoma gástrico;
- Analisar a expressão diferencial de genes humanos entre amostras positivas e negativas para a presença do vírus Epstein -Barr;
- Identificar a expressão dos genes do vírus Epstein-Barr em amostras positivas;

3 REVISÃO LITERÁRIA

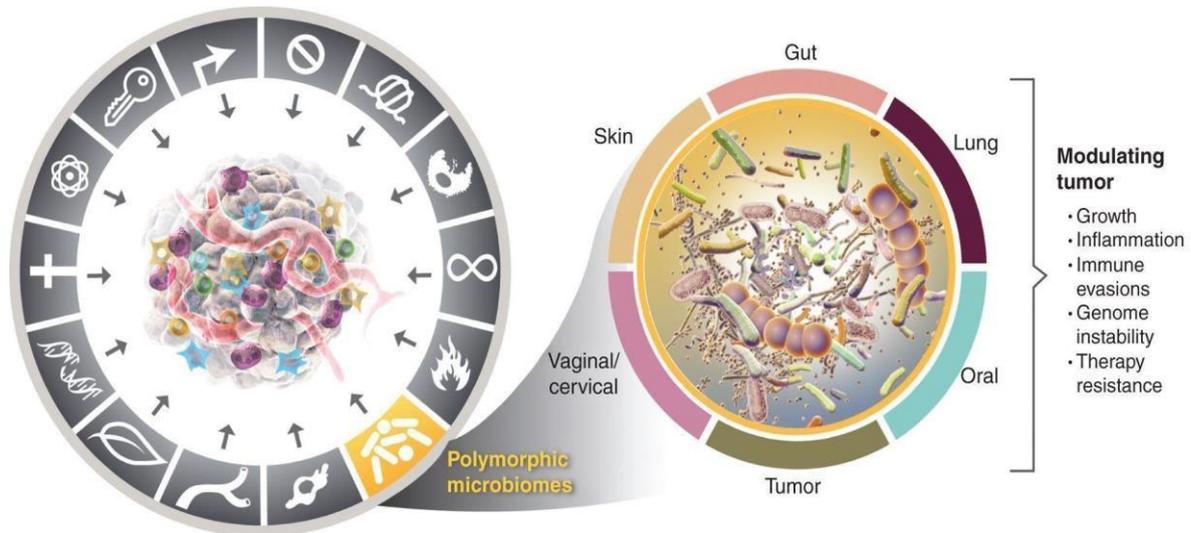
3.1 Câncer Gástrico

O câncer surge como resultado de disfunções nos mecanismos regulatórios que conferem vantagens seletivas às células cancerosas. As células tumorais apresentam diferenças significativas em relação às células normais, devido a falhas nos sistemas reguladores, resultando em diversas alterações genéticas, tais como mutações nas sequências de genes, variações no número de cópias, inserções, deleções e eventos de recombinação (Grady et al., 2020).

Em 2000, Douglas Hanahan e Robert Weinberg propuseram que o processo neoplásico necessitava de 6 capacidades distintas que podem ser identificadas para compreender a complexidade da doença, sendo divididas entre características funcionais e de habilitação. Ao longo dos anos e com a evolução nas pesquisas, hoje, as capacidades, chamadas de “*hallmarks of cancer*”, apresentam 14 marcos. Alguns dos marcos estudados consistem em: inflamação promotora de tumor, indução de angiogênese, resistência à morte celular, ativação de invasão e de metástase, reprogramação epigenética não mutacional, senescência celular, instabilidade e mutação do genoma e microbiomas polimórficos (Hanahan; Weinberg, 2000; Hanahan, 2022).

Os microbiomas polimórficos podem influenciar - com efeitos deletérios ou protetivos - no desenvolvimento tumoral, na progressão maligna e/ou na resposta à terapia. Hanahan (2022) também sugere que uma manifestação desses microrganismos pode ser a criação de microambientes imunológicos promotores ou antagonistas de tumores, que podem proteger ou facilitar a tumorigênese e a progressão tumoral (figura 1).

Figura 1 - Hallmarks of Cancer, com foco nos polimorfismos microbianos.



Fonte: Hanahan 2022.

Existem diversas abordagens de classificação do CG baseadas tanto em características histológicas quanto em perfis moleculares. Uma das classificações histológicas mais utilizadas é a proposta por Laurén, em 1965, que divide o CG em dois subtipos distintos: o subtipo intestinal e o subtipo difuso (Lauren, 1965; Waldum; Fossmark, 2018).

O subtipo intestinal, o qual é mais prevalente, é caracterizado pela presença de uma estrutura glandular e células com morfologia mais definida e variável. Quanto às suas origens, esse subtipo muitas vezes está associado à presença da bactéria *H. pylori*, uma vez que a inflamação gástrica provocada por essa bactéria pode resultar em uma série de lesões pré-neoplásicas (Lauren, 1965; Mera et al, 2017; Correa; Piazzuelo, 2012).

Por outro lado, o subtipo difuso é menos comum e se caracteriza por uma aparência mais difusa. Esse tipo de tumor tende a se disseminar mais amplamente dentro da mucosa gástrica e está associado a fatores genéticos, como mutações em proto-oncogenes (Waldum; Fossmark, 2018). As células cancerosas desse subtipo podem induzir fibrose, levando a uma redução na elasticidade e complacência do tecido, o que pode resultar em sintomas semelhantes aos observados no câncer gástrico cirrótico. O subtipo difuso se diferencia do subtipo intestinal em incidência, infecção por *H. pylori* e possui um prognóstico menos favorável, devido às suas características menos definidas (Assumpção, 2020; Lauren, 1965).

As classificações por perfis moleculares no CG estão divididas em 4 subtipos distintos. O primeiro é o subtipo de tumores com instabilidade de microssatélites (MSI), que abrange aproximadamente 21,7% dos tumores analisados. O segundo é o subtipo estável genomicamente (GE), representando cerca de 19% dos casos. O terceiro subtipo é caracterizado

pela instabilidade cromossômica (CIN), o qual abrange a maioria das neoplasias, aproximadamente 49,8%. Por fim, há o subtipo positivo para o vírus Epstein-Barr (EBVaCG), que compreende cerca de 8,8% das ocorrências de câncer gástrico. Os subtipos moleculares mostram características genômicas distintas, as quais podem ser um guia para a compreensão da heterogeneidade tumoral (Cancer Genome Atlas Research Network, 2014).

Os tumores MSI apresentam uma notável taxa de mutação, inclusive em genes encarregados da codificação de proteínas essenciais para a sinalização oncogênica. A manifestação desse fenômeno é primordialmente atribuída à metilação do gene promotor, culminando no silenciamento transcricional do gene *MLH1*. Este último é responsável pelo processo de reparo de incompatibilidades no DNA. Adicionalmente, essa subcategoria de tumores demonstra uma incidência significativa na expressão do gene *PD-L1*, sendo correlacionada com a presença de *H. pylori* e metaplasia intestinal (Wang et al., 2019; Cancer Atlas Genome Research Network, 2014).

O subtipo tumoral GE apresenta como principal característica a alta expressão em vias relacionadas a angiogênese, bem como mutações no gene *CDH1* e no gene *RHOA*, que faz parte da família homóloga Ras. Essas alterações moleculares são correlacionadas com o aumento de padrões de crescimento heterogêneos, além de contribuírem para a invasão e a perda de coesão celular, características intrínsecas aos tumores difusos (Wang et al., 2019; Cancer Atlas Genome Research Network, 2014).

No subtipo neoplásico CIN, suas características são delineadas pela presença de elevadas taxas de mutações no gene *TP53*, indicando uma ativação focal da via do receptor tirosina quinase RAS (RTKS/RAS). Além disso, evidencia-se a amplificação dos genes das ciclinas E1 (CCNE1) e D1 (CCND1), os quais desempenham um papel crucial na regulação do ciclo celular. Essas amplificações não apenas influenciam o ciclo celular, mas também abrem caminho para possíveis alvos terapêuticos (Wang et al., 2019; Cancer Atlas Genome Research Network, 2014).

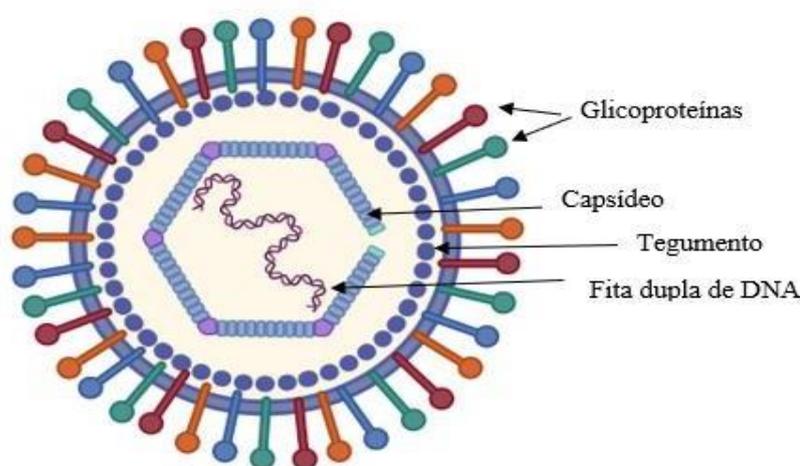
O EBVaCG descreve a forte associação entre o vírus com a carcinogênese gástrica. Suas características estão relacionadas a alterações epigenéticas como a hipermetilação acentuada do DNA, mutações nos genes *PIK3CA*, *BCOR* e *ARID1A*. Além dessas alterações, observa-se a amplificação do gene *JAK2* e uma superexpressão das moléculas *PD-L1* e *PD-L2*. Um traço adicional deste subtipo é o silenciamento do gene *CDKN2A*, responsável por codificar duas proteínas que desempenham um papel crucial nas redes de supressão de tumores (Sunakawa; Lenz, 2015; Cancer Atlas Genome Research Network, 2014).

Com base nas análises realizadas pelo TCGA, a classificação molecular do CG contribui significativamente para a condução de estudos mais abrangentes em busca de potenciais biomarcadores. Isso engloba elementos como os anticorpos *PD-1/PD-L1* e os inibidores de *PI3K* para os casos de EBVaCG. Dessa maneira, essas descobertas têm o potencial de desempenhar um papel crucial na melhoria do tratamento de pacientes com câncer, ao introduzir novas abordagens para alcançar resultados clínicos mais favoráveis. Isso também abre caminho para a implementação de terapias personalizadas e direcionadas (Sunakawa; Lenz, 2015; Cancer Atlas Genome Research Network, 2014).

3.2 Human gammaherpesvirus 4 (Epstein-Barr)

O *Human gammaherpesvirus 4* ou Epstein-Barr Vírus, pertencente à família *Herpesviridae*, foi o primeiro vírus a ser relacionado a patologias malignas. Similarmente a alguns agentes virais, previamente investigados na literatura, o EBV se configura como um vírus envelopado, incorporado em um capsídeo icosaédrico, que protege seu material genético, além de apresentar um tegumento que protege o capsídeo (figura 2). Adicionalmente, seu genoma totaliza aproximadamente 179 kb e engloba a codificação de aproximadamente 100 produtos gênicos, os quais demonstraram correlação com seus mecanismos de replicação (Shinozaki-Ushiku et al., 2014; Zavala-Vega et al., 2019).

Figura 2 - Estrutura do vírus Epstein-Barr.



Fonte: Adaptado de Zavala-vega *et al.* (2019).

Devido à sua atividade oncogênica, o genoma do EBV pode ser identificado em diversas neoplasias, tais como câncer nasofaríngeo, linfoma de Hodgkin, linfoma de Burkitt e CG. A

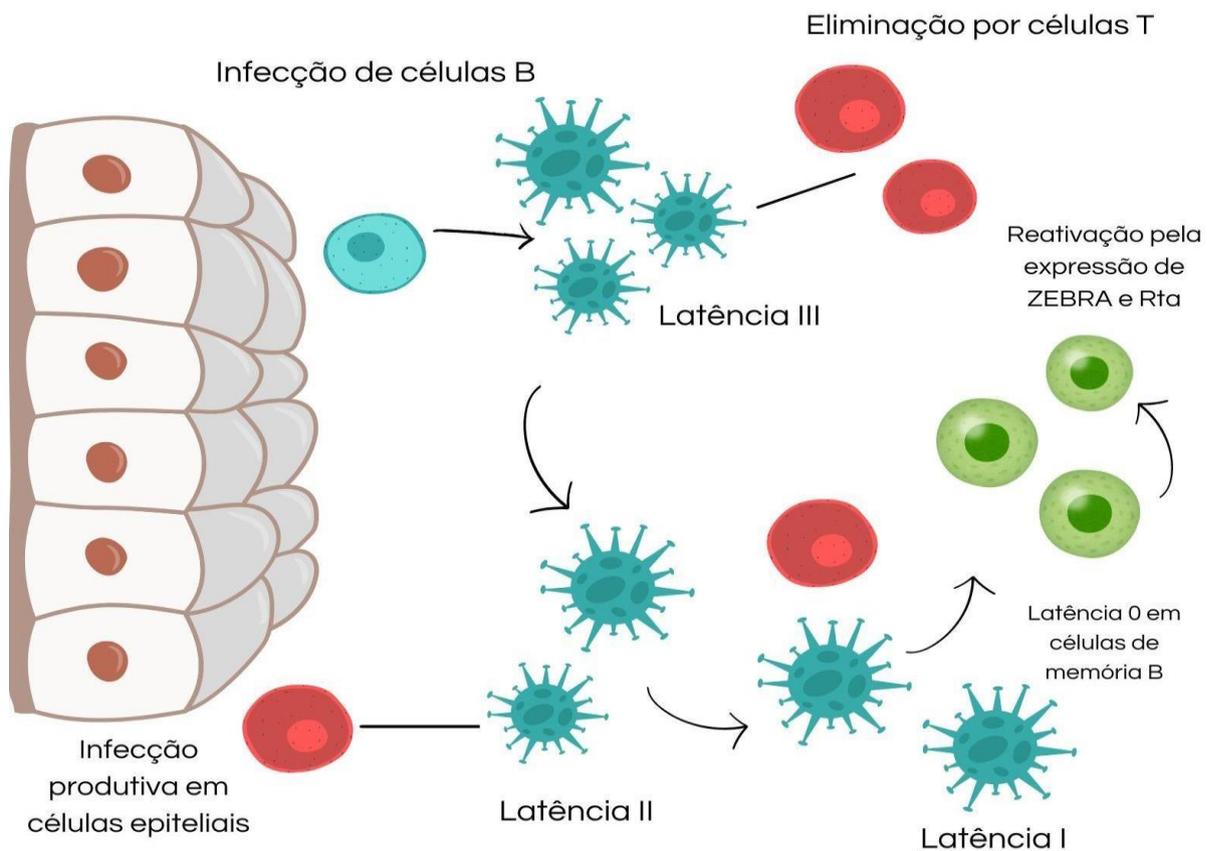
detecção da presença do vírus em pacientes com CG ocorreu inicialmente em 1990, por meio de análises utilizando técnicas de PCR. No entanto, somente com a aplicação das técnicas de *hibridização in situ* para a identificação de pequenos RNAs codificados pelo EBV (EBERs) foi possível aprimorar a identificação do vírus em amostras de CG (Iizasa et al., 2012; Shinozaki-Ushiku et al., 2014).

A literatura ainda mostra que a presença do patógeno dispõe de características clinicopatológicas e aberrações genômicas únicas. O EBV associado ao câncer gástrico (EBVaCG) possui uma taxa de incidência estimada entre 75.000 - 90.000 novos casos anuais. Além disso, a prevalência de EBVaCG é mais alta em pacientes do sexo masculino, com menos de 60 anos, diagnosticados com tumores do tipo difuso ou tumores na região proximal. No entanto, pacientes com esse subtipo molecular apresentam melhor taxa de sobrevida global quando comparados com aqueles que são diagnosticados com tumores negativos para EBV (Hirabayashi et al., 2023; Huang et al., 2014).

3.2.1 Replicação viral

A infecção pelo vírus ocorre por meio da via oral, predominantemente afetando linfócitos B e células epiteliais. Inicialmente, a replicação do vírus se inicia no local da infecção, especificamente nas células epiteliais da orofaringe, permitindo assim a disseminação para as células B do sistema imunológico. A entrada do EBV nas células B apresenta maior eficiência devido à sua interação com um receptor de alta afinidade, o CD21, o qual se liga à glicoproteína gp350 do vírus (figura 3). Além disso, o vírus desenvolve um estado de latência, utilizando proteínas latentes como o *EBNA-1*, que impede o sistema imunológico de eliminá-lo (Zavala-Vega et al., 2019).

Figura 3 - Ciclo de replicação do EBV.



Fonte: Adaptada de Zavala-Vega *et al.*, 2019.

O EBV apresenta dois tipos de infecções: a infecção lítica e a infecção latente. Na maior parte do tempo, o vírus permanece em estado de infecção latente, caracterizado pela ausência de replicação. (Borozan *et al.*, 2018).

O ciclo lítico é onde ocorre a replicação do genoma viral, produção de novas cópias do patógeno e propagação do vírus para novas células hospedeiras. Dessa forma, os genes líticos produzidos nesse ciclo são subdivididos em genes precoces imediatos, precoces e tardios, responsáveis pela manutenção da fase lítica. A indução da replicação se inicia com a expressão de *BZLF1* e *BRLF1* - genes precoces imediatos - que atuam diretamente na replicação do DNA, além de participar na ativação e expressão de genes que ajudam o vírus na resposta imune do hospedeiro (Zhang; Xu, 2023).

Os genes precoces são expressamente induzidos por meio dos genes precoces imediatos citados acima e estão relacionados a replicação do DNA viral e codificam enzimas envolvidas no metabolismo de nucleotídeos. Enquanto os genes tardios são responsáveis pela codificação de componentes estruturais para a manutenção do DNA do vírus. Após estabelecer uma

replicação primária nas células epiteliais, o vírus se propaga pelas células B e inicia uma infecção latente (Zhang; Xu, 2023).

A infecção latente é caracterizada por quatro formas distintas, as quais são determinadas pelos padrões de expressão dos genes latentes, variando de acordo com o estado fisiológico e patológico do hospedeiro. A expressão gênica em cada tipo de latência resulta na ativação de genes associados a diferentes tipos de doenças malignas (quadro 1) (Sun *et al.*, 2020).

A primeira fase de latência que se estabelece nas células B infectadas é do tipo III. No entanto, essas células são reconhecidas e eliminadas pelo sistema imunológico, especialmente pelas células T. Além disso, o EBV pode direcionar diretamente as células de memória e estabelecer uma latência do tipo 0, que não apresenta expressão significativa de genes associados a patologias (Zavala-Vega *et al.*, 2019).

Segundo a análise realizada por Strong e colaboradores (2013), quando o vírus está associado ao CG, ele se enquadra em uma fase intermediária entre as latências do tipo I ou II. Tal fato ocorre mesmo que haja uma expressão limitada de proteínas específicas desses tipos, muitas vezes em uma manifestação epigenética de baixa intensidade. A infecção por EBV pode ter consequências significativas, visto que além da atividade oncogênica, o vírus também possui atividade imunogênica (Strong *et al.*, 2013; Wong *et al.*, 2022).

Quadro 1 – Expressão gênica do ciclo de latência do EBV.

Fases de latência	Genes expressos	Alterações malignas	Patologia associada
0	EBER's	Nenhuma	-
I	EBNA 1, EBER's & BART's	Linfócitos Quiescentes; Ausência de resposta de linfócitos T citotóxicos	Linfoma de Burkitt; EBV associado ao CG
II	LMP1 & LMP2A		Linfoma de Hodgkin; Linfoma de células NK; Carcinoma nasofaríngeo
III	EBNA-LP, EBNA-2 & EBNA-3	Proliferação celular e ativação das respostas de linfócitos T citotóxicos	Doença linfoproliferativa pós-transplante

Fonte: Adaptado de Zavala-Vega *et al.* (2019).

3.2.1 Alterações genéticas em EBVaCG

Uma alteração significativa e característica do EBVaCG é a hipermetilação do DNA em promotores de genes, que desempenham um papel crucial na supressão de tumores e na regulação de processos como o ciclo celular, adesão celular, metástases, apoptose e reparo de DNA. Essa hipermetilação pode ocorrer tanto no genoma do hospedeiro quanto no genoma do vírus. A metilação de ilhas CpG em regiões promotoras de genes supressores de tumor precede o silenciamento da transcrição, sendo um dos principais mecanismos envolvidos na carcinogênese (Usui *et al.*, 2021).

Fatores como o envelhecimento e o uso de cigarros podem influenciar os níveis de metilação do DNA em ilhas CpG ao longo da vida. Além disso, algumas proteínas virais, como *EBNA1* e *LMP2A*, podem desencadear a hipermetilação por meio da superexpressão da DNA metiltransferase. No entanto, de acordo com Shinozaki-Ushiku e colaboradores (2015), a metilação de genes virais pode ser um mecanismo de defesa do hospedeiro para suprimir a infecção pelo EBV (Shinozaki-Ushiku *et al.*, 2015; Yang *et al.*, 2020).

As expressões virais do Epstein-Barr mais incidentes no CG são de proteínas do ciclo latente como: *EBER's*, *EBNA1*, *BART*, *BARF1* e em alguns casos há a expressão de *LMPI* e *LMP2A*. Os *EBER's* (*EBER-1* e *EBER-2*) são pequenos RNAs não codificantes transcritos pelo EBV utilizando a enzima RNA polimerase III. A expressão dos *EBER's* está intrinsecamente relacionada a vários processos celulares, incluindo a proliferação celular, resistência à apoptose e interações com fatores de crescimento e proteínas do hospedeiro. Em particular, o *EBER-1* desempenha um papel fundamental na modificação da expressão de miRNAs, com o objetivo de suprimir o gene *CDH1* (Shinozaki-Ushiku *et al.*, 2015).

A proteína viral *EBNA-1* é expressa em todos os tipos de malignidades associadas à infecção pelo vírus. Sua presença é crucial para a replicação e estabilização dos epissomas do EBV. Além disso, o *EBNA-1* desempenha um papel relevante na desregulação epigenética de genes específicos para o CG, como as gastrocinas 1 e 2, que são considerados genes supressores de tumor nesse contexto. Além das suas funções na replicação viral e na modulação epigenética dos genes supressores de tumor, o *EBNA-1* apresenta potencial oncogênico. No entanto, é importante ressaltar que a expressão do *EBNA-1* em casos de câncer gástrico é frequentemente reduzida ou até mesmo inexistente, o que representa um desafio significativo na sua utilização como alvo terapêutico específico para essa condição (Yang *et al.*, 2020).

Os transcritos *BART* são expressos em grande parte das doenças causadas pelo EBV e incluem genes como *BARF0*, *BARF1*, *RPMS1* e *A73*, que codificam proteínas específicas. O *BARF1* é um dos genes que se manifesta em EBVaCG e está associado à resistência a tratamentos quimioterápicos e ao silenciamento de genes relacionados à proliferação e à apoptose. Para promover a proliferação de células de câncer gástrico, o *BARF1* regula positivamente a via de sinalização NF-κB/ciclina D1, que desempenha um papel fundamental no controle do crescimento celular (Shinozaki-Ushiku et al., 2015; Yang et al., 2020). Essas descobertas ressaltam a importância do *BARF1* como um componente crítico na patogênese do EBVaCG e como um alvo potencial para estratégias terapêuticas.

Além disso, os *BARTs* regulam as proteínas *LMP1* e *LMP2A* e essa regulação negativa desempenha um papel fundamental no controle da latência e reativação do vírus. Em particular, a proteína *LMP1* possui diversas funções, incluindo a promoção da proliferação, migração celular e indução de alterações epigenéticas no genoma do hospedeiro.

As mutações mais frequentemente encontradas em tumores positivos para o EBV envolvem os genes *PIK3CA*, *ARID1A* e *BCOR*, juntamente com a amplificação de *JAK2*, *PD-L1* e *PD-L2*. Quando o gene *PIK3CA* está mutado, ele sofre uma amplificação anormal, contribuindo para a proliferação e sobrevivência celular na tumorigênese gástrica. Isso ocorre através da ativação da via de sinalização PI3K/Akt, que é regulada nas células neoplásicas quando há perda da função da proteína *PTEN* e ganho de função das proteínas *PIK3* em virtude das mutações no gene *PIK3CA*, responsável pela subunidade catalítica (Sunakawa; Lenz, 2015; Yang et al., 2020).

Nesse sentido, se faz necessária a identificação molecular dos mecanismos de progressão do ciclo celular e oncogênese causados pelo EBV. Essa abordagem é essencial para aprimorar tanto o diagnóstico quanto às opções terapêuticas empregadas na prática clínica. Além disso, a identificação de marcadores moleculares pode contribuir para o diagnóstico precoce, a estratificação de pacientes com base em riscos e características individuais, bem como o acompanhamento mais preciso da progressão da doença (Yin et al., 2019).

3.3 Tratamentos e Diagnósticos

Conforme o exposto, a pesquisa e o desenvolvimento de terapias mais direcionadas e eficazes para cânceres associados ao EBV continuam sendo uma área proeminente a se explorar, assim, estratégias como a imunoterapia e a terapia-alvo estão sendo buscadas como

alternativas promissoras para melhorar os resultados e a sobrevida dos pacientes com EBVaCG (Wong *et al.*, 2022).

O tratamento de cânceres positivos para o EBV envolve abordagens terapêuticas clássicas, como cirurgia, radioterapia e/ou quimioterapia. A escolha do tratamento específico depende do tipo e estágio do câncer, bem como das características individuais do paciente. No entanto, é importante reconhecer que as respostas aos tratamentos podem variar consideravelmente. Vários tipos de câncer positivos para o EBV apresentam altas taxas de recidiva, o que significa que a doença pode retornar após um período de remissão. Além disso, alguns desses cânceres podem ser altamente agressivos e desenvolver resistência aos tratamentos convencionais, o que torna o manejo da doença desafiador (Wong *et al.*, 2022; Yang *et al.*, 2020).

A detecção de EBVaCG pode ser feita por meio de testes imuno-histoquímicos, *hibridização in situ* (ISH), PCR digital e por sequenciamento do genoma. Os testes imunohistoquímicos (IHC) são focados na proteína *LMPI* codificada pelo EBV, entretanto não consegue detectar a localização ou quantidade transcricional do vírus. A *hibridização in situ* é feita por uma sonda que detecta especificamente o *EBER-1* em amostras de tecido fixadas com formaldeído e parafina, entretanto este método ocasionalmente pode resultar em falsos positivos e falsos negativos. Quando utilizados em conjunto, o ISH e o teste imuno-histoquímicos podem melhorar a precisão da detecção e reduzir as chances de resultados errôneos (Sun *et al.*, 2020).

O método de triagem baseado em PCR digital começou a ser empregado em 2019, sendo utilizado para calcular a carga de DNA do vírus baseado no número de cópias dos fragmentos BamH1-W do EBV (Shuto *et al.*, 2019).

Além disso, em 2015, Camargo e colaboradores testaram ferramentas de sequenciamento de nova geração (NGS) para realizar a detecção do vírus em amostras de CG. As amostras escolhidas foram amostras com status EBV positivo por método ISH que, então, foram submetidas à análise de *whole-genome*, *whole-exome*, *mRNA* e *miRNA*. Das 23 amostras que possuíam status EBVaCG, apenas 1 caso não foi validado em ambos os métodos. Nesse sentido, esse estudo mostrou que utilizar dados de NGS pode ser uma alternativa para os testes convencionais de detecção de EBVaCG e que, para além disso, a disponibilidade desses dados pode auxiliar em descobertas e melhores entendimentos a respeito dos perfis moleculares de cada paciente (Camargo *et al.*, 2016).

3.4 Metatranscriptoma por RNA-Seq

Os métodos de análise a partir de dados de RNA-Seq estão desempenhando um papel importante tanto nas pesquisas clínicas quanto na biologia moderna. Além disso, com os dados de RNA-Seq é possível a investigação quantitativa de todo o transcriptoma de forma altamente eficiente e abrangente (Deshpande *et al.*, 2023).

O transcriptoma é definido como a totalidade das transcrições de genes presentes em uma célula durante um estágio de desenvolvimento particular ou em uma condição fisiológica específica e permite a identificação de genes diferencialmente expressos, vias de sinalização envolvidas e a compreensão dos mecanismos subjacentes a esses processos (Wang *et al.*, 2009).

Nesse contexto, o estudo do metatranscriptoma tem como objetivo principal a descrição da composição taxonômica das comunidades microbianas e sua capacidade funcional. Além disso, a metatranscriptoma busca aprofundar a compreensão das interações entre os microrganismos e seus hospedeiros, a identificação de vias metabólicas ativas de relevância, bem como a avaliação de como as funções expressas podem influenciar a progressão de doenças (Shakya *et al.*, 2019).

4 METODOLOGIA

4.1 Aspectos éticos e obtenção das amostras

O projeto foi submetido ao Comitê de Ética e Pesquisa (CEP) do HUIBB e aprovado sob número 47580121.9.0000.5634, conforme publicado na Plataforma Brasil.

Para este estudo, foram utilizadas 76 amostras de tecido tumoral de pacientes diagnosticados com AG submetidos a ressecção cirúrgica. Os pacientes convidados a participar do estudo foram advertidos a respeito dos benefícios, malefícios e possíveis riscos em participar desta pesquisa, e assim, concordaram em participar de forma voluntária e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) (Anexo I).

4.2 Extração de RNA total

Foram macerados cerca de 50-100 ng de tecido tumoral de cada amostra para a extração de RNA, realizada através do método de TRIzol Reagent (Thermo Fisher Scientific), seguindo as instruções do fabricante. Em seguida, o RNA total foi avaliado quanto à integridade e concentração (ng/μL), utilizando-se o equipamento Qubit 4 Fluorometer (Thermo Fisher Scientific). Os critérios estabelecidos para o RIN (número de integridade de RNA) ≥ 5 .

4.3 Construção das bibliotecas e sequenciamento de RNA

A construção das bibliotecas foi realizada com o kit *TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Gold* (Illumina, US) de acordo com as instruções do fabricante, utilizando cerca de 1 μg de RNA total por amostra, em um volume total de 11 μL. Após a finalização das bibliotecas, as amostras foram submetidas a uma avaliação de integridade dos fragmentos de DNA gerados, realizada no equipamento 2200 TapeStation System (Agilent Technologies AG, Basel, Suíça).

O sequenciamento foi executado em *pair-end* na plataforma NextSeq® 500 (Illumina®, US), utilizando o kit *NextSeq® 500 MID Output V2 kit - 150 cycles* (Illumina®) de acordo com as instruções do fabricante.

4.4 Controle de qualidade das leituras

Após o sequenciamento, as leituras obtidas de RNA-Seq foram submetidas ao controle de qualidade. Inicialmente, a qualidade das leituras foi avaliada utilizando o FastQC. Posteriormente, os adaptadores e as leituras de baixa qualidade foram removidas usando o *Trimmomatic v0.39*, no qual adotou-se como parâmetro um valor de qualidade (QV) superior a 15.

4.5 Classificação Molecular de EBVaCG

Para a realização da classificação molecular de EBVaCG, as leituras foram alinhadas com a ferramenta *Kraken2* (Wood; Lu, 2019), utilizando como index o *Reference Viral Database (RVDB)* (Goodacre *et al.*, 2018), que consiste em um banco de dados com maior especificidade para identificação de sequências virais em dados genômicos e transcriptômicos.

4.6 Alinhamento das amostras com o genoma humano

As leituras de transcritos humanos foram caracterizadas através de alinhamento e quantificação com o *Salmon v1.5.2*, utilizando como index de referência os transcritos codificantes no hg v38 (www.ensembl.org) e utilizando como anotação o *GENCODE v.42* (www.gencodegenes.org). As leituras foram importadas do *Salmon* para o RStudio com o pacote *Tximport v3.14.0*.

4.7 Análises Multivariadas

4.7.1. Análise de Expressão Diferencial de Genes Humanos

Após a classificação e o alinhamento, foi realizada uma análise de expressão gênica diferencial entre amostras tumorais EBVaCG e EBVnCG (CG não associado a EBV), utilizando o pacote *DESeq2* (Love; Huber; Anders, 2014). Os genes considerados diferencialmente expressos (DE) foram aqueles cuja diferença de expressão alcançaram os critérios: i) $|\text{Log}_2(\text{Fold-Change})| > 1$; e ii) $p\text{-ajustado} < 0.05$.

Os resultados da expressão gênica diferencial foram visualizados utilizando os pacotes *ggplot2* v3.5.1 (Wickham, 2016) e *complexHeatmap* v2.14.0 (Zuguang, 2017).

4.7.2. Análises da Característica de Operação do Receptor

Para determinar a especificidade e a sensibilidade dos genes diferencialmente expressos, foi calculada a área sob a curva (AUC) da característica de operação do receptor (ROC) com a utilização do pacote *pROC* v1.18 (Robin *et al.*, 2017). Foram selecionados os genes com AUC > 0.85.

4.7.3. Análises de enriquecimento funcional dos genes

Os genes DE com AUC > 0.85 foram enriquecidos por meio da Ontologia Genética para termos relacionados a processos biológicos, com o intuito de compreender melhor a sua relevância biológica e funcional. O pacote *Cluster Profile* v4.8.3 (Guangchuang, 2017) foi utilizado para essa análise.

4.8 Análise de Expressão de Genes Virais

Para avaliar a expressão dos genes virais do EBV, as amostras que foram consideradas positivas de acordo com o método de RNA-Seq foram selecionadas e submetidas ao pipeline *nf-core/rnaseq* v3.14 (Patel *et al.*, 2024). Foram utilizados como index o genoma de referência do EBV GCF_002402265.1 e a anotação genômica NC_007605.1, ambos disponíveis no banco de dados do *National Center for Biotechnology Information* (NCBI).

4.8.1 Correlação da expressão dos genes humanos DE com genes virais

Para avaliar a correlação da expressão gênica dos pacientes com os genes virais expressos, foi realizada a correlação de *Spearman* onde 1 significa correlação positiva e -1 significa correlação negativa. Em seguida, para a visualização gráfica das correlações, foi utilizado o pacote *corrplot* no Rstudio.

5 RESULTADOS

5.1 Classificação Molecular do EBV

Para a classificação molecular de EBVaCG, o critério estabelecido para considerar uma amostra como positiva foi a detecção de um número de leituras de genes virais ≥ 100 . Os resultados obtidos pelo método de RNA-Seq foram comparados com os dados de teste imunohistoquímico (IHC), realizado pelo departamento de patologia do HUIBB, para avaliar a paridade de resultados (quadro 2).

Das 76 amostras tumorais selecionadas, 8 amostras foram consideradas positivas para a presença do vírus pelos dois métodos. Dentre as amostras testadas pelo teste IHC, 1 amostra não demonstrou leituras suficientes para ser considerada positiva pelo método de RNA-Seq. As demais amostras consideradas negativas por IHC apresentaram leituras dos genes virais ≤ 100 (Anexo II).

Quadro 2 - Comparação das amostras positivas entre método de RNA-Seq e IHC.

AMOSTRAS	RNA-Seq	IHC
170111	11435	POSITIVA
190708	51919	POSITIVA
190807	25456	POSITIVA
191204	7951	POSITIVA
200330TU	42172	POSITIVA
210209T	404	POSITIVA
220829T	4472	POSITIVA
180830	58808	POSITIVA
190724	1	POSITIVA

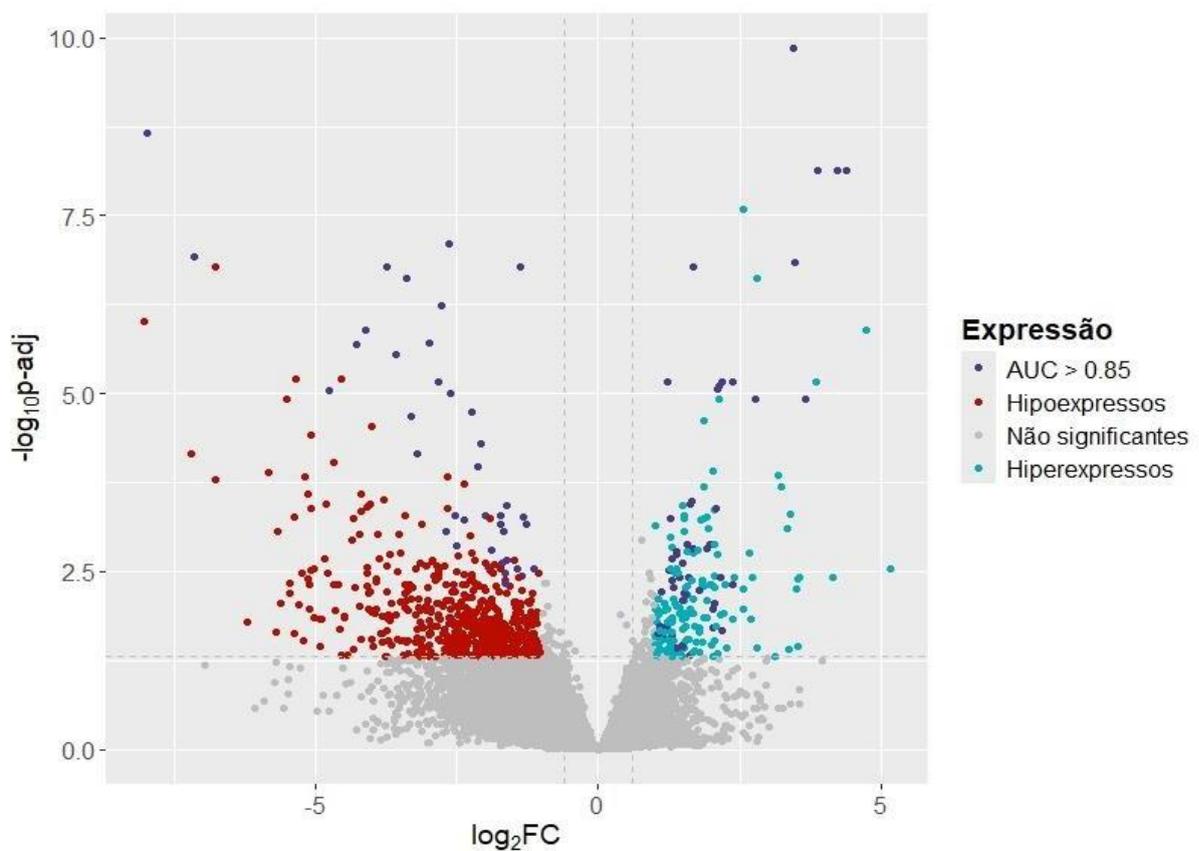
Fonte: Elaborada pelo próprio autor.

5.2 Análise da Expressão Gênica Diferencial

Após a classificação das amostras, foi realizada uma análise do perfil de expressão gênica diferencial entre os pacientes EBVaCG e EBVnCG. Nessa análise, 834 genes humanos foram diferencialmente expressos entre as amostras de acordo com os critérios estabelecidos para genes DE. Dentre esses genes, 612 foram hipoexpressos, enquanto 222 foram hiperexpressos.

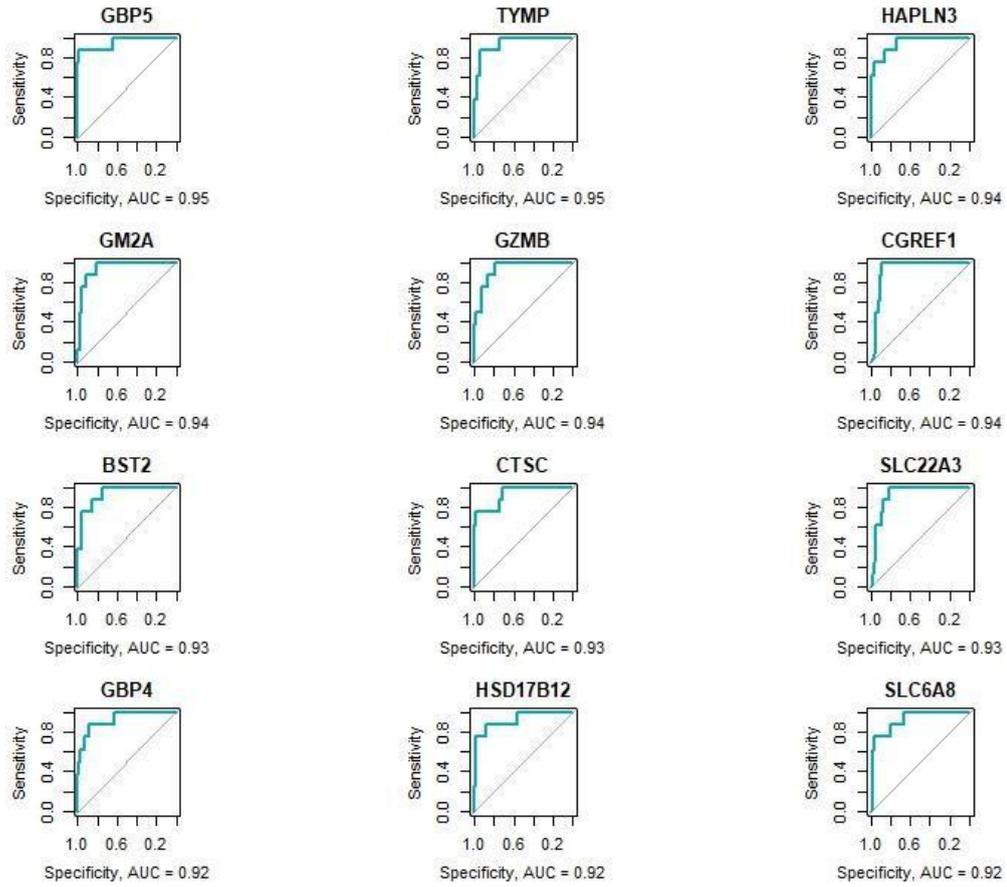
Para avaliar quais genes DE possuem maior especificidade e sensibilidade na discriminação entre as amostras EBVaCG e EBVnCG, selecionamos aqueles que apresentaram o valor de AUC > 0.85 (figura 4). Dentre os 834 genes DE, 92 foram selecionados através do valor de corte estabelecido pela AUC, nos quais selecionamos os 20 com maiores valores para investigar a influência na carcinogênese (Tabela 1). Os genes *GBP5* e *TYMP* se destacam pois apresentaram AUC > 0,95 (figura 5).

Figura 4 - *Volcano plot* exibindo os genes DE entre amostras positivas e negativas para a presença de EBV.



Fonte: Elaborada pelo autor.

Figura 5 - Gráficos de curva ROC representando a AUC por gene.



Fonte: Elaborado pelo autor.

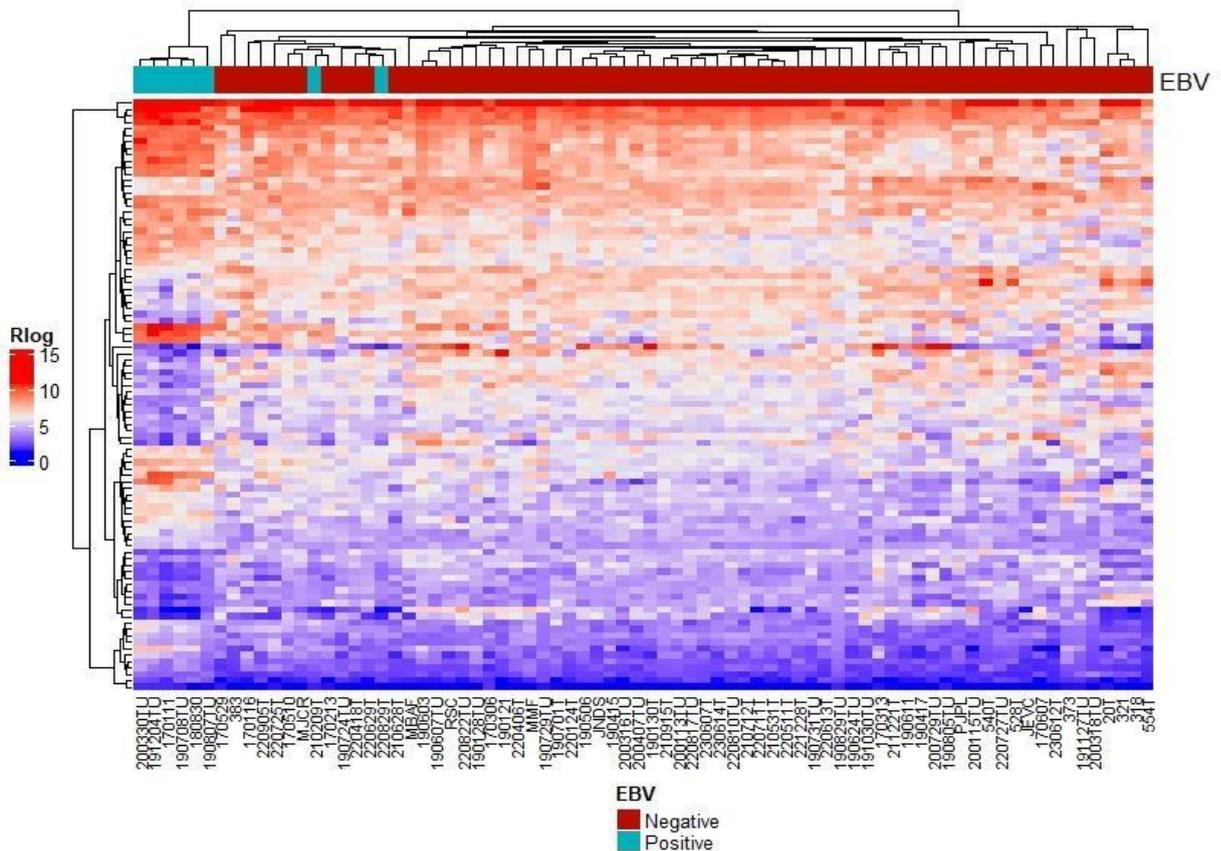
Tabela 1 - Genes humanos DE com AUC > 0.85.

GENES	AUC	log2FC	p-adj
GBP5	0,95	3,442	0,000000000142
TYMP	0,95	2,111	0,000008814488
HAPLN3	0,94	2,379	0,000006749687
GM2A	0,94	1,224	0,000006749687
GZMB	0,94	2,768	0,000012197105
CGREF1	0,94	-3,741	0,000000164985
BST2	0,93	2,096	0,000402177807
CTSC	0,93	1,672	0,000000164985
SLC22A3	0,93	-3,401	0,00000024445
GBP4	0,92	2,193	0,000006749687
HSD17B12	0,92	-1,379	0,000000164985
SLC6A8	0,92	-2,999	0,000001965821
WNK2	0,92	-2,819	0,000006749687
CXCL9	0,91	3,472	0,000000141881
CAMK2N1	0,91	-2,776	0,000000582084
UBE2L6	0,91	1,399	0,001847402087
NSMF	0,91	-2,367	0,000609625692
PCLO	0,91	-4,106	0,000001262813
ZNF683	0,91	2,369	0,004770528182
PPP1R1B	0,91	-4,276	0,000002022704

Fonte: Elaborado pelo autor

Para uma melhor visualização dos padrões de expressão gênica entre as amostras EBVaCG e EBVnCG, selecionamos apenas os genes com $AUC > 0,85$ e conduzimos uma análise de agrupamento hierárquico não supervisionado (figura 6). Conseguimos observar que as amostras EBVaCG apresentam um padrão de expressão gênica distinto das amostras EBVnCG, o que pode-se ser percebido na formação de *clusters*, que demonstram que determinados grupos de genes exibem regulação negativa em amostras EBVaCG do que amostras EBVnCG. Além disso, o *heatmap* demonstra que as amostras foram agrupadas de acordo com suas semelhanças nos padrões de expressão.

Figura 6 - *Heatmap* dos padrões de expressão dos genes DE com $AUC > 0,85$.



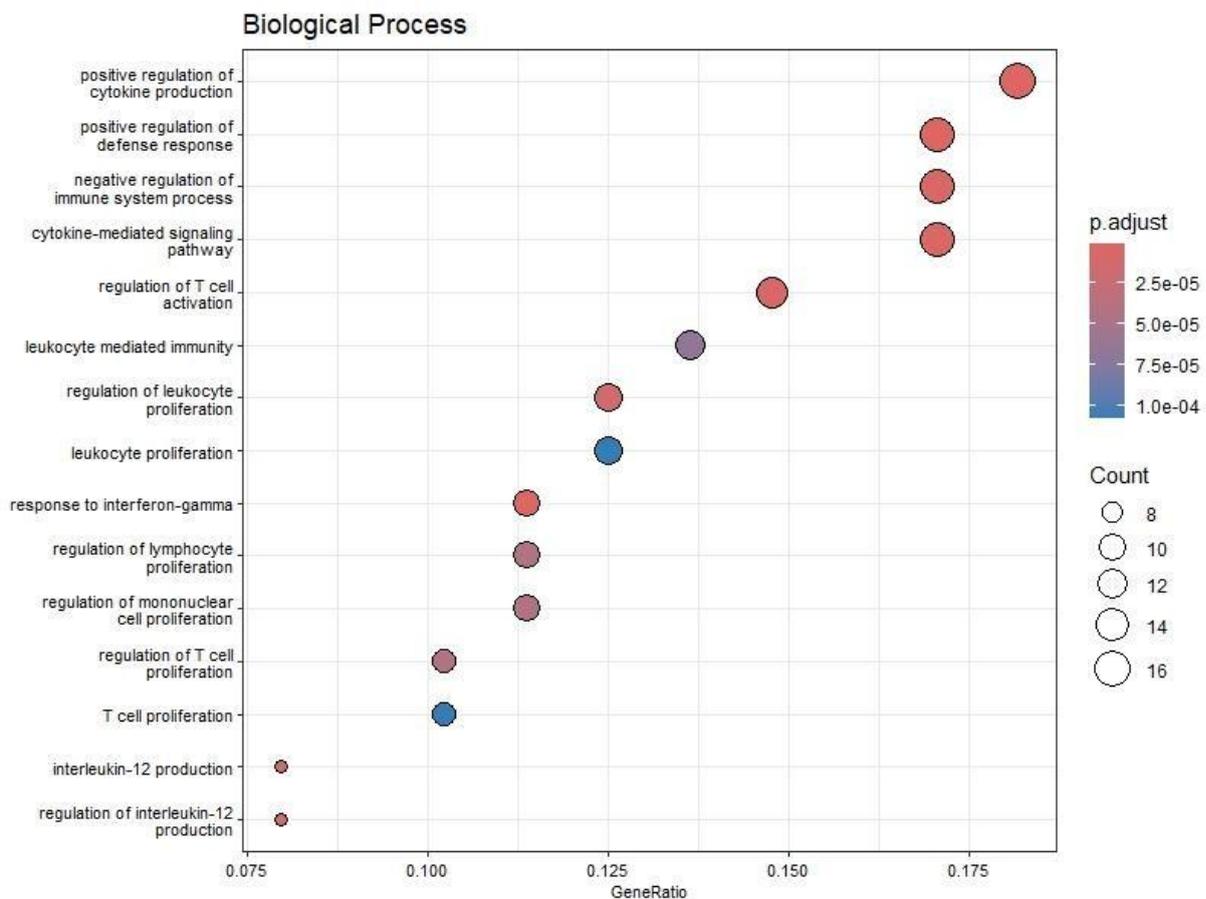
Fonte: Elaborada pelo autor.

5.2.1 Análise de enriquecimento funcional dos genes DE

Para maior entendimento das funções e influências dos genes DE realizamos a ontologia pelo *Gene Ontology (GO)* disponível no pacote *ClusterProfile*, nos quais obtivemos 316 termos

enriquecidos significativos para processos biológicos. Dentre os principais termos enriquecidos, podemos destacar alguns relacionado a resposta imune, tais como: a regulação da ativação da célula T, regulação positiva da resposta de defesa, regulação da proliferação de células T, imunidade mediada por leucócitos e regulação da proliferação de linfócitos (figura 7).

Figura 7 - Ontologia Gênica para processos biológicos.



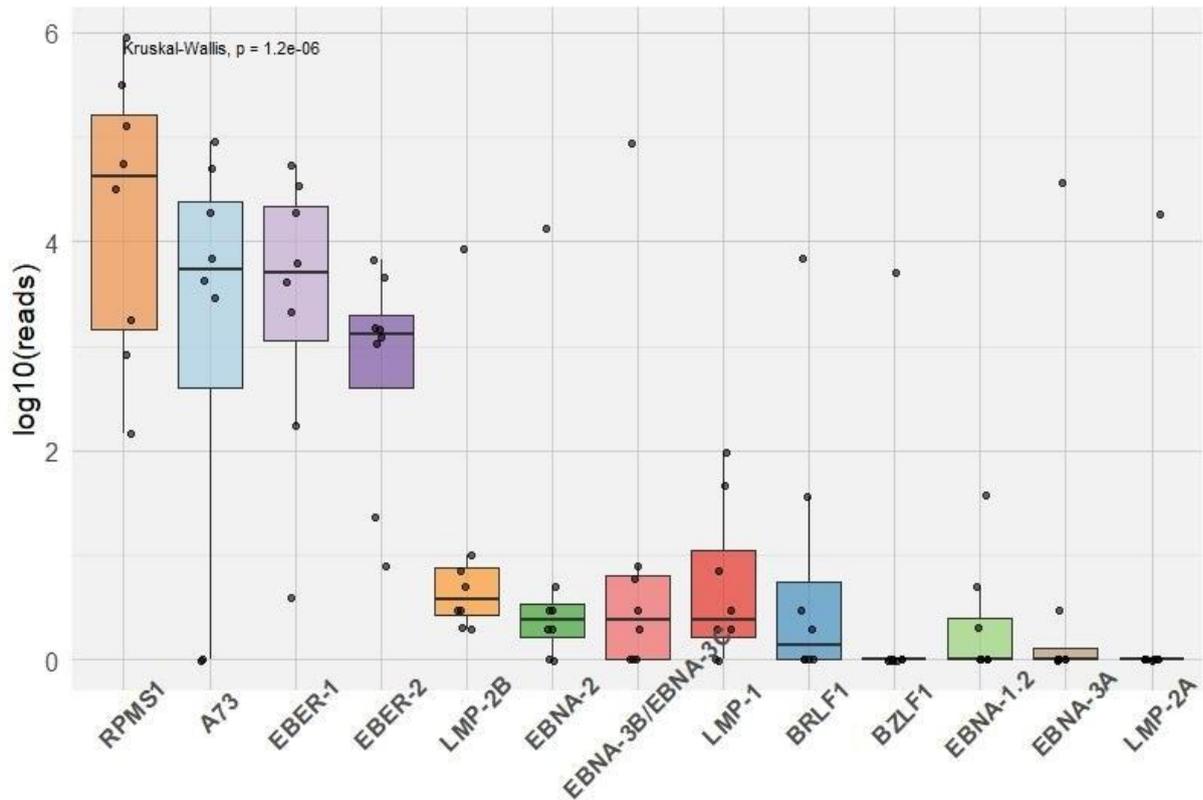
Fonte: Elaborado pelo autor.

5.3 Análise da Expressão de Genes Virais

Nossas análises foram direcionadas para avaliar a expressão de genes virais nas amostras positivas para EBV. Através do alinhamento utilizando o pipeline do *nf-core RNAseq*, obtivemos os valores dos *counts*, que consistem na quantificação de quantas vezes um gene ou transcrito específico é representado nas leituras sequenciadas.

Dentre as expressões mais frequentes, identificamos os genes *RPMS1*, *A73* e *LMP-2B* presentes em todas as amostras. Igualmente, constatamos alta expressão de *EBER-1* e *EBER-2* que são pequenos RNAs não codificantes importantes para a identificação do vírus através de testes IHC (figura 8).

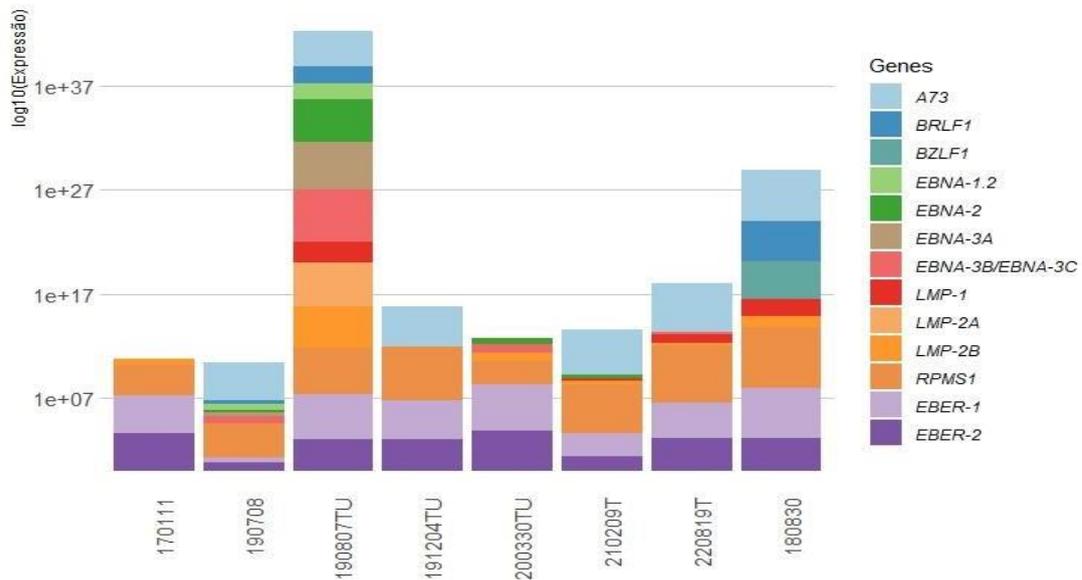
Figura 8 - Expressão dos genes virais de EBV.



Fonte: Elaborado pelo autor.

Em seguida, avaliamos as expressões gênicas virais por amostras. A amostra 190807 é a que demonstra maior expressão dos genes de EBV, enquanto a amostra 170111 possui expressão de apenas 4 genes virais (figura 9).

Figura 9 - Expressão dos genes virais do EBV por amostras.



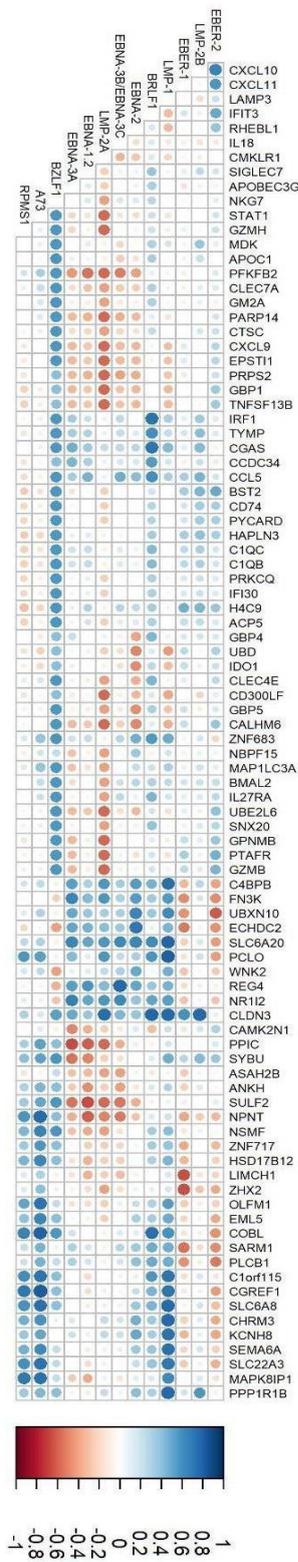
Fonte: Elaborado pelo autor.

5.3.1 Correlação da expressão dos genes humanos DE com genes virais

Após analisar a expressão dos genes virais nas amostras positivas, correlacionamos esses genes com os genes DE humanos previamente analisados (figura 10). Em azul, é representada uma correlação positiva entre os genes, enquanto que em vermelho mostra uma correlação negativa. Conseguimos identificar a formação de 2 clusters maiores, o primeiro com a correlação negativa entre os genes virais e genes DE e o segundo com correlação positiva.

O primeiro cluster é composto pelos seguintes genes humanos: *PFKFB2*, *CLEC7A*, *GM2A*, *PARP14*, *CTSC*, *CXCL9*, *EPST11*, *PRPS2*, *GBP1*, *TNFSF13B*. Esses genes encontram-se correlacionados negativamente aos genes virais da família *EBNA* (*EBNA-1.2*, *EBNA-2*, *EBNA-3A*, *EBNA-3B/EBNA-3C*) e a *LMP-2A*. O segundo cluster é formado pelos genes humanos: *C4BPB*, *FN3K*, *UBXN10*, *ECHDC2*, *SLC6A20*, *PCLO*, *WNK2*, *REG4*, *NR1I2* e *CLDN3*, que por sua vez encontram-se correlacionados positivamente com os genes virais da família *EBNA* (*EBNA-1.2*, *EBNA-2*, *EBNA-3A*, *EBNA-3B/EBNA-3C*), com dois genes da família *LMP*, a *LMP-1A* e a *LMP-2A* e por fim com um gene da fase lítica, o *BRLF1*.

Figura 10 - Correlação entre os genes DE com AUC > 0,85 e os genes virais expressos nas amostras positivas.



Fonte: Elaborado pelo autor.

6 DISCUSSÃO

A identificação da presença do EBV em amostras de CG possui considerável relevância clínica, pois, além das características moleculares e histológicas específicas, pacientes com diagnóstico positivo para essa infecção viral demandam abordagens terapêuticas distintas. Nesse sentido, buscamos testar um método de classificação a partir de dados de RNA-Seq, inferir sobre a expressão gênica diferencial dos pacientes com EBVaCG e EBVnCG e investigar quais são as principais expressões virais nas amostras positivas para a presença do vírus.

De acordo com Murphy e colaboradores (2009), a taxa de prevalência do EBV em tumores gástricos é de aproximadamente 9%, dessa forma a classificação molecular do EBVaCG se mostra extremamente importante para uma compreensão mais completa da interação entre o vírus e a expressão gênica humana, e pode fornecer informações valiosas para o desenvolvimento de estratégias terapêuticas direcionadas e aprofundar nosso conhecimento sobre as bases moleculares das doenças associadas à presença do EBV.

Na literatura há diversos estudos que já foram conduzidos sobre a presença do vírus em amostras de câncer gástrico, entretanto os métodos utilizados para a classificação molecular e identificação de genes virais diferem do que o nosso estudo propõe. Camargo e colaboradores (2016) e TCGA (2014) utilizaram métodos como *whole genome*, *whole exome*, sequenciamento de *mRNA e miRNA*, além de validar por meio de testes IHC e ISH. Em contrapartida, realizamos análises de dados provenientes de RNA-Seq e testes IHC para a confirmação de nossos achados.

Nossos resultados mostram que das 76 amostras analisadas, 8 apresentam status positivo para a presença do vírus através do método de RNA-Seq, o que corresponde a aproximadamente 10,52% dos tumores gástricos estudados, corroborando com o que há descrito na literatura acerca da prevalência do EBV em tumores gástricos (Camargo *et al.*, 2016; Tavakoli *et al.*, 2020).

Foi realizada uma análise de expressão gênica diferencial dos genes humanos entre as amostras positivas e negativas para EBV. Aproximadamente 834 genes foram considerados DE, e 92 deles apresentaram $AUC > 0.85$. Dos 92, selecionamos 20 genes com os maiores valores de AUC, para inferirmos a respeito de sua relação com a carcinogênese e/ou com o EBV e se poderiam ser considerados biomarcadores de EBVaCG.

O gene *GBP5* é descrito como um gene atuante na ativação do inflamossoma *NLRP3*, que são complexos que alertam o sistema imunológico a respeito de infecção ou algum dano tecidual, possuindo assim um papel importante na imunidade inata (Shenoy *et al.*, 2012). Na

literatura, há um estudo comparando genes DE dos perfis moleculares de amostras EBVaCG contra amostras de CG negativo para EBV (EBVnCG) que demonstrou a hiperexpressão deste gene em amostras positivas (Wang *et al.*, 2021). Em nossos achados, corroborando a literatura existente, observamos a hiperexpressão do gene *GBP5* em amostras positivas para a presença do vírus, com uma AUC superior a 0,95.

Outro gene encontrado hiperexpresso em nossas amostras positivas foi o *GBP4*, que é um importante gene da família *GBP*. O gene indica maior inflamação no microambiente tumoral, o que pode explicar sua alta expressão em amostras positivas para EBVaCG visto que essa é uma das características importantes do EBVaCG (Shinozaki-Ushiku, Kunita, Fukayama, 2015). Além disso, estudos demonstram que o *GBP4* pode estar fortemente correlacionado a fatores imunológicos em diversos tipos de câncer e também pode se correlacionar positivamente com o gene *PD-L1/CD274* (Zhou *et al.*, 2024).

O gene *CD274* é relatado com frequência em vários tipos de cânceres humanos. A ligação de PD-L1 ao receptor PD-1 nas células T ativadas inibe a imunidade antitumoral, neutralizando os sinais que ativam essas células. Este gene possui um importante papel quando tratamos de EBVaCG, visto que sua hiperexpressão pode contribuir para a evasão imunológica, dificultando a detecção e eliminação das células malignas pelo sistema imunológico. Além disso, ele é considerado um importante alvo no tratamento para cânceres associados ao EBV através da imunoterapia (Sun; Mezzadra; Schumacher, 2018). O gene *CD274* também foi encontrado hiperexpresso em nossas amostras EBVaCG.

Seguindo com nossas análises, o gene *TYMP* também revelou uma regulação positiva em amostras EBVaCG, apresentando uma AUC > 0.94. Não encontramos descrição na literatura que o relacione com o EBV ou com o subtipo molecular EBVaCG. Em contrapartida, há associações do gene com CG, onde a expressão de *TYMP* pode influenciar na carga de mutação do tumor. Além disso, o mesmo estudo observa uma correlação positiva entre a expressão do gene e a presença de células T CD8+ (Yang *et al.*, 2022).

Pode-se inferir que quando hiperexpresso em amostras positivas para EBV, o gene *TYMP* pode aumentar a capacidade de infiltração em células T CD8+ visto que o perfil molecular de amostras EBVaCG frequentemente exhibe infiltração significativa de células T CD8+ em comparação ao perfil de amostras EBVnCG (Salnikov *et al.*, 2023). Portanto, faz-se necessário investigar a hiperexpressão do gene *TYMP* em amostras com perfis EBVaCG e verificar quais suas influências no microambiente tumoral.

Apesar de estar altamente expresso em nossas amostras positivas, não encontramos relação com EBV ou EBVaCG com o gene *HAPLN3*. Há diversos estudos que relacionam este gene a diversos outros tipos de câncer. Dentre eles, encontramos um que faz associação da hipermetilação do promotor deste gene em amostras de câncer de próstata (Møller, *et al.*, 2017). Além disso, também há associações do gene com proliferação, migração e invasão celular em células de câncer renal (Ding *et al.*, 2023). Entretanto, não encontramos estudos relacionando a expressão deste gene a CG.

Em nossas buscas, alguns genes não se mostraram direta ou indiretamente associados a EBVaCG, EBV ou CG. Dentre eles está o gene *GM2A*, ainda que tenha apresentado hiperexpressão em amostras positivas, não há estudos de sua relação com câncer e está associado a doenças degenerativas. O gene *CGREF1*, por outro lado, já foi descrito como um gene importante para o prognóstico de câncer colorretal e também codifica proteína de nome semelhante, que está envolvida na regulação do crescimento e da proliferação celular. Ele pode atuar como um fator de controle negativo, inibindo a divisão celular em resposta a certos sinais (Deng *et al.*, 2015; Xiang *et al.*, 2023). Em nossas amostras EBVaCG, o *CGREF1* apresentou hipoexpressão.

O *BST2*, assim como no presente trabalho, já demonstrou regulação positiva em amostras de CG. Para além disso, foi demonstrado que sua hiperexpressão pode estar relacionada à promoção da transição epitélio-mesenchimal (TEM) e também a metástase em CG, quando induzido por *HOXD9* (Li *et al.*, 2024). Entretanto, não encontramos qualquer relação com EBVaCG ou EBV.

Igualmente, o gene *CTSC* fora relacionado com a promoção de migração celular, habilidades de invasão celular e também com a transição epitélio-mesenchimal quando se trata de células de glioma (Li *et al.*, 2024). Além disso, também há um estudo onde este gene está diretamente ligado ao câncer de mama, no qual ele promove a metástase da mama para o pulmão através da regulação de neutrófilos (Xiao *et al.*, 2021). Todavia não há descrito na literatura qualquer estudo de sua expressão em EBVaCG ou CG.

A hiperexpressão do *GZMB* em CG foi descrita como uma influência regulatória na biologia celular do CG, visto que pode estar associada a promoção de crescimento e migração celular, além da EMT (Lu *et al.*, 2024). Em nossos resultados o gene se encontra hiperexpresso em amostras positivas para EBVaCG, o que pode indicar uma resposta imunológica ativa contra as células tumorais visto que o *GZMB* é uma protease envolvida na indução de apoptose de

células-alvo quando liberada por linfócitos citotóxicos ou células *natural killer* (NK) (Hagn *et al.*, 2015).

O gene *SLC22A3* foi hipoexpresso em amostras positivas para EBVaCG. Na literatura há a descrição da superexpressão desse gene em linhagens celulares de câncer colorretal, a qual pode estimular migração e invasão celular (Cervenkova *et al.*, 2019). Além disso, há outros estudos associando a diversos outros tipos de câncer todavia não foram observadas associações do gene com o vírus ou até mesmo com CG. Com efeito, o gene *HSD17B12*, também regulado negativamente em amostras positivas, não possui relação com EBVaCG ou CG descrita até o momento. Em 2011, foi publicado um estudo que sugere sua participação em vias críticas que conectam o processo de inflamação a câncer, mas que ainda precisam de maiores investigações para tal confirmação (Visus *et al.*, 2011).

Os genes *SLC6A8*, encontram-se hipoexpressos em nossas amostras e não encontramos descrição que façam associações ao subtipo molecular EBVaCG ou a CG. Há pesquisas sugerindo que sua função como transportador de creatina 1, tem possibilidade de desempenhar um papel importante no metabolismo energético das células cancerígenas influenciando na progressão tumoral (Kurth *et al.*, 2021).

O *WNK2* foi associado como um supressor tumoral em estudos de câncer hepatocelular, além do registro do silenciamento epigenético em outros tipos de cânceres (Zhou *et al.*, 2019). Também o identificamos hipoexpresso em nossos resultados. Podemos inferir que sua expressão baixa pode favorecer no progresso tumoral, visto que Du e colaboradores (2020) demonstraram que quando este gene encontra-se metilado por um rna longo não codificante, o *LINC00858*, há a restrição da apoptose e aumento da invasão e migração celular em células de CG (Du *et al.*, 2020).

Seguindo nossas análises, detectamos que o gene *CAMK2N1* encontra-se com níveis de expressão baixo nas amostras positivas. O subtipo molecular EBVaCG tem como uma das características o bom prognóstico de seus pacientes. Desse modo, podemos explicar a hipoexpressão do gene como um biomarcador para o prognóstico uma vez que estudos mostram que a alta expressão deste gene, quando mediada por RNAs não codificantes, está relacionada a um mau prognóstico e a migração celular em CG (Peng; Ren; Ren, 2022).

O gene *NSMF*, também hipoexpresso, ainda não foi descrito com associações a EBVaCG ou a CG, embora, quando desregulado, possa causar instabilidade cromossômica. Além disso, sua função está atrelada a rede de respostas ao estresse da replicação celular para manutenção do genoma (Ju *et al.*, 2021). O gene *PCLO* já foi associado a função neural, e

quando hiperexpresso em CG pode atrasar a progressão tumoral (Bernhardt *et al.*, 2024). Em nossas amostras, ele encontra-se hipoexpresso, podendo assim diminuir seu controle em vias de sinalização. Para confirmar esta hipótese, são necessários maiores estudos sobre sua influência na carcinogênese associada ao EBV.

Encontramos *ZNF683* regulado positivamente em nossas amostras. Este gene foi descrito como um fator de transcrição que tem como papel fundamental a diferenciação de células imunes NK (Post *et al.*, 2017). Dessa forma, sugere-se que sua hiperexpressão pode influenciar na eficiência da resposta imunológica em tumores gástricos.

Em nossos resultados, detectamos a hiperexpressão do *CXCL9* nas amostras positivas, reforçando o que já é descrito na literatura sobre a família de genes *CXC* estar mais presente em EBVaCG do que EBVnCG (Mu *et al.*, 2022). Podemos inferir que a alta expressão deste gene pode estar influenciando diretamente na hiperexpressão do gene já citado *CD274/PD-L1*, visto que a expressão do *CXCL9* pode regular a expressão de *PD-L1* por vias de sinalização, tais como STAT e PI3K-Akt (Zhang *et al.*, 2018).

Outro gene com alta expressão em nossas amostras foi o *UBE2L6*. Não há pesquisas descrevendo correlação deste gene com EBVaCG ou EBV. Porém, há um estudo que propõe que o *UBE2L6* pode ser utilizado como biomarcador de prognóstico para CG (Wang *et al.*, 2023).

Por fim, o gene *PPP1R1B* demonstrou regulação negativa em nossas amostras. Na literatura, encontramos apenas associação com CG no qual o gene se liga ao eixo de sinalização *IGF1R-SRC* para promover a ativação da via STAT3, que é capaz de promover a proliferação e invasão de células tumorais, além de induzir quimiorresistência (Ashrafizadeh *et al.*, 2020; Tiwari *et al.*, 2020). Também não há associações com EBVaCG ou EBV.

Os resultados discutidos até aqui sugerem um papel fundamental dos genes DE no desenvolvimento e progressão tumoral. Para melhor compreender a função destes genes na biologia do CG, foi realizada uma ontologia genética, a qual revelou que a maioria dos genes DE estão associados a processos biológicos do sistema imune e de defesa do organismo.

A regulação positiva da produção de citocinas é um processo importante para o sistema imunológico e envolve a participação de vários genes que codificam as proteínas necessárias para a sinalização celular e para a resposta inflamatória. Além disso, as citocinas podem desempenhar papéis importantes no desenvolvimento do câncer e também controlando os diferentes estágios da doença por meio de interleucinas, fatores estimuladores de colônias, fatores de crescimento transformadores, interferons e quimiocinas (Yasmin *et al.*, 2015).

Uma citocina importante e hiperexpressa em nossos genes DE é o *IFNG*. Essa regulação positiva pode levar a um aumento na produção de outras citocinas, uma vez que, é um dos genes envolvido na regulação da produção de citocinas e, também, na regulação positiva da resposta de defesa. Sua função como gene pode contribuir para a resposta imune inata ao reprogramar macrófagos para o fenótipo pró-inflamatório M1 e células T promovendo a produção de citocinas adicionais que podem potencializar a resposta imune contra células tumorais (Jorgovanovic *et al.*, 2020).

Sabemos que a resposta imune dispõe de regulações positivas e regulações negativas, sendo ambas importantes para o controle inflamatório e para a prevenção de doenças autoimunes. Nesse sentido, nossa ontologia gênica mostrou que há genes DE envolvidos no processo de regulação negativa do processo do sistema imune, dentre os quais podemos citar *PDCD1LG2/PD-L2*. O *PD-L2* é um ligante do receptor *PD-1*, uma proteína expressa em células T e que está associada à exaustão imunológica, um dos mecanismos da regulação negativa do sistema imune. A exaustão imunológica pode ser caracterizada pela perda das funções das células T, podendo levar até a apoptose dessas células (Vigano *et al.*, 2012).

Outro processo biológico contemplado na ontologia gênica é a regulação da ativação das células T, que constitui em um processo intrincado que abarca fatores de transcrição, diversos sinais de sinalização e mecanismos regulatórios (Salmond, 2023). No contexto do CG, as células T reguladoras (Tregs) desempenham um papel crucial na patogênese e na manutenção do equilíbrio central entre supressão imunológica e autoimunidade. Além disso, as Tregs também são importantes no que diz respeito a regulação da proliferação de células T, visto que possuem a capacidade de limitar a expressão das células T auxiliares e das células T citotóxicas CD8+, promovendo efeitos supressores (Wang *et al.*, 2023).

Os processos biológicos de imunidade mediada por leucócitos, regulação da proliferação de leucócitos e proliferação de leucócitos são essenciais para o sistema imunológico adaptativo e inato. Os leucócitos são divididos por tipos e funções, dentre eles estão os linfócitos, especificamente as células T e B. A imunidade mediada por leucócitos pode ocorrer através de mecanismos, tais como: reconhecimento de antígenos, produção e ativação de citocinas e quimiocinas, que são fundamentais para a comunicação entre células imunes e para o recrutamento de leucócitos para os locais de inflamação ou tumor (Chi; Pepper; Thomas, 2024).

A presença de termos de ontologia relacionados à resposta imune adaptativa em nossas amostras indica que a presença do EBV no microambiente tumoral pode induzir uma forte

resposta imune antiviral. Esta resposta envolve a reativação das células de defesa, que podem combater tanto o vírus quanto as células tumorais. Além disso, já foi registrado que o EBV utiliza diversas estratégias para evadir a resposta imune, tanto inata quanto adaptativa. No entanto, em certos contextos, especialmente nas infecções latentes, o sistema imunológico pode ser capaz de detectar e responder ao vírus, mobilizando células T CD8+ e outras respostas imunes adaptativas para controlar a infecção e potencialmente atacar as células tumorais associadas ao vírus (Silva; Alves; Pontes, 2024).

O EBV codifica diversas proteínas, nos quais os padrões de codificação associados a diversos tipos de malignidade são variáveis. Além disso, a infecção latente do vírus pode ser dividida em quatro fases e cada fase expressa diferentes genes virais. O EBVaCG já foi correlacionado a apenas uma fase da infecção, a fase I, entretanto estudos recentes o relacionam também à fase II dada a expressão de genes dessa fase (Zavala- Vega *et al.*, 2019). Nesse sentido, nossas análises seguiram para avaliar quais genes virais estavam mais expressos em nossas amostras positivas e quais influências esses genes podem ter na oncogênese do CG.

De acordo com nossos achados, *EBER-1* e *EBER-2* encontram-se expressos em 100% das amostras positivas, confirmando o que há descrito na literatura. Esses produtos virais são pequenos RNA's não codificantes e encontram-se expressos em todas as fases da infecção latente, além de serem utilizados como marcadores diagnósticos para a detecção do EBV (Bai *et al.*, 2022).

Podemos associar a expressão dos *EBERs* diretamente a um aumento da resposta imunológica visto que, Kondo e colaboradores afirmam que a perda de expressão do gene viral *EBER* está associada a diminuição da expressão do *PD-L1* e de linfócitos T CD8+, ou seja, ao regular negativamente ou diminuir os antígenos relacionados ao vírus, o tumor pode escapar da vigilância imunológica, o que beneficia a progressão tumoral (Kondo *et al.*, 2023).

Os transcritos *BART* são descritos como uma combinação de microRNAs (miRNAs) e longos RNAs não codificantes (lncRNAs) expressos durante a fase latente da infecção. Suas funções são associadas a evasão imunológica e promoção da carcinogênese (Shinozaki-Ushiku *et al.*, 2015). Dentre nossos resultados, encontramos expressos alguns transcritos como: *RPMS1* e *A73*.

Há um estudo que sugere que o *RPMS1* é expresso apenas durante circunstâncias específicas e pode contribuir para o desenvolvimento do câncer, ou seja, que a expressão deste transcrito seja encontrada em um estágio pré-canceroso mas não em células cancerígenas (Jang; Jung; Kim, 2011). O *A73* ao ser testado *in vitro* demonstrou a capacidade de ter funções

bioquímicas que possam ser relevantes para a carcinogênese (Morale-Sánchez; Fuentes-Panana, 2017).

Diferindo do que está descrito na literatura onde a expressão desses genes não foram encontrados em células cancerígenas *in vivo*, nossos resultados mostram a manifestação desses produtos virais. Os genes *RPMS1* e *A73* foram detectados em 100% e 75% das amostras, respectivamente, sugerindo que esses genes podem estar envolvidos não só no estágio pré-câncer mas também contribuindo para a manutenção da carcinogênese.

Além dos mencionados acima, encontramos expressos genes pertencentes à fase lítica da infecção, chamados de genes de início imediatos ou *immediate-early genes* (IE). *BRLF1* e *BZLF1* encontram-se manifestados em 25% e 12,5% das amostras, respectivamente. Já foi demonstrado que *BZLF1* pode alterar a expressão de vários genes hospedeiros e que também pode atuar suprimindo respostas do sistema inflamatório a fim de viabilizar a replicação do vírus. Enquanto que o *BRLF1* pode regular tanto a replicação do DNA viral quanto a transcrição do mRNA (Long *et al.*, 2021). Apesar de não haver estudos associando esses genes a EBVaCG, a compreensão destes e suas funções podem nos fornecer informações importantes para o desenvolvimento de estratégias terapêuticas contra a infecção pelo EBV.

Encontramos também a expressão de genes virais da família *LMP*, são eles *LMP-1*, *LMP-2A* e *LMP-2B*. A manifestação destes em nossas amostras pode ser explicado pelo fato das *LMP* apresentarem propriedades oncogênicas e potencial para transformar células epiteliais, ativando as vias de sinalização e modificando a expressão de oncogenes e supressores tumorais (Li *et al.*, 2016).

Em nossos resultados, a *LMP-1* está expressa em 4 amostras positivas, ou seja, 50%. Entretanto, os níveis de expressão são baixos, corroborando com o que há descrito na literatura onde se discute que apesar de ser um importante oncogene, a expressão de *LMP-1* em CG é baixa ou inexistente devido a fase de latência em que se expressa (Kang; Kieff, 2015). Além disso, a baixa expressão deste gene pode ser explicada por um estudo que sugere que o mesmo pode ter sido silenciado através da metilação que o seu promotor tende a sofrer (Li *et al.*, 2016).

Por outro lado, a *LMP-2A*, encontrada em 53,8% dos casos de EBVaCG descritos na literatura, apresenta expressão em apenas 12,5% de nossas amostras. Uma das ações da *LMP-2A* é a ativação da fosforilação da proteína *STAT3*, que pode se ligar à enzima *DNA metiltransferase 1* (DNMT1). Quando essa ligação é regulada positivamente, resulta no silenciamento da expressão de supressores tumorais por meio da metilação das ilhas CpG em

sua região promotora. Essas mudanças epigenéticas, especialmente a hipermetilação do DNA no genoma do hospedeiro, representam um dos principais mecanismos envolvidos na indução da carcinogênese do EBVaCG (Hino *et al.*, 2009; Tavakoli *et al.*, 2020).

Em uma revisão sistemática a respeito da expressão dos genes virais, Ribeiro e colaboradores (2017) publicaram que a manifestação da *LMP-2B* acontece em apenas 10% dos casos. Nossos resultados diferem do que está relatado na literatura, apresentando expressão da *LMP-2B* em 75% das nossas amostras. Com a baixa expressão desse gene em casos de EBVaCG, também há a falta de estudos sobre sua função no desenvolvimento do CG. Há estudos descrevendo que a *LMP-2B* possui o potencial de regular a função da *LMP-2A*, prevenindo a potencial lise do EBV, enquanto altos níveis de expressão da *LMP-2B* também podem acelerar a transição da infecção latente para a proliferativa do EBV (Tang *et al.*, 2015).

Por fim, há a expressão de genes da família *EBNA* em nossos resultados. Dentre esses resultados encontramos *EBNA 1.2*, o qual podemos considerar uma variante do *EBNA1*, com expressão baixa em 25% de nossas amostras. Este gene é considerado essencial para a replicação do DNA viral e manutenção do epissoma durante a replicação celular. Além disso, regula a expressão de outros genes do EBV, controlando a transcrição de genes virais importantes para a latência (El-Sharkawy; Al Zaidan; Malki, 2018).

O gene *EBNA-2*, foi encontrado expresso em 50% de nossos resultados. A expressão deste gene acontece logo após a infecção das células B pelo vírus. Em um estudo publicado em 2023, por Zhang e colaboradores, há a sugestão de que a expressão desse gene juntamente com a expressão de *LMP1* podem alterar a transformação de células B primárias em linhagens celulares linfoblásticas (Zhang *et al.*, 2023). Não há descrição na literatura da presença de *EBNA-2* no subtipo EBVaCG, visto que normalmente encontra-se expresso na fase 3 da latência e o subtipo EBVaCG manifesta genes das fases 1 ou 2.

Outros genes em que a expressão em EBVaCG não é relatada são os pertencentes a *EBNA-3*. Entretanto, em nossas amostras há a expressão de *EBNA-3A* em 25%, que apresentou um nível de expressão alto em apenas uma amostra, e também há manifestação de *EBNA-3B/EBNA-3C* em 50%. Esses genes funcionam como reguladores da transcrição do vírus e da célula hospedeira. Além disso interagem com fatores de transcrição para transativar ou reprimir a expressão gênica. Sua manifestação está relacionada a fase 3 da latência (El-Sharkawy; Al Zaidan; Malki, 2018).

Em síntese, nossos achados confirmam a expressão de vários genes virais do EBV nas amostras positivas, alinhando-se com descrições prévias na literatura, ao mesmo tempo que

revelam um novo ponto de vista sobre a presença e potencial impacto desses genes nas células cancerígenas. A expressão de *EBER-1* e *EBER-2*, assim como a presença dos transcritos *BART*, *RPMS1* e *A73*, sugerem uma contribuição contínua para a manutenção da carcinogênese. Enquanto que a identificação de genes da fase lítica, como *BRLF1* e *BZLF1*, em algumas amostras, juntamente com a manifestação das *LMPs* e *EBNAs*, ressalta a complexidade da infecção latente do EBV e a diversidade dos seus padrões de expressão gênica no CG.

Em vista disso, buscamos correlacionar os genes virais e suas expressões com os genes humanos diferencialmente expressos. Não encontramos estudos que descrevam diretamente a correlação dos genes DE com os genes virais, então, se tornou necessário entender em quais mecanismos os genes presentes nos clusters identificados estão associados.

Dentre os genes presentes no cluster 1 encontramos alguns que estão associados a vias biológicas importantes, tais como: imunidade, inflamação e apresentação de antígenos. Na via de imunidade e inflamação, os genes relacionados são *CLEC7A*, *CXCL9* e *TNFSF13B* (Daley *et al.*, 2017; Mao *et al.*, 2021; Tokunaga *et al.*, 2018) envolvidos na ativação e sobrevivência de células imunes. Nesse sentido, podemos sugerir que a expressão de genes virais como *EBNA-2* e *LMP-2A* podem influenciar negativamente na regulação da expressão desses genes visto que a expressão desses produtos virais podem alterar a transformação de células B (Zhang *et al.*, 2023).

Além dos mencionados, também podemos sugerir que o gene *PFKFB2* envolvido na regulação do metabolismo celular, pode ser influenciado pela atividade de fatores de transcrição ativado por genes virais como *EBNA-2* (Ozcan *et al.*, 2020; Zhang *et al.*, 2023).

No cluster 2 encontramos pelo menos 10 genes DE que estão correlacionados positivamente com os genes virais, no entanto, não há descrições que expliquem essas correlações. Podemos inferir sobre alguns genes como o *PCLO*, que está associado à sinapse e pode influenciar a comunicação celular, além de ativar células imunes, o que é fundamental na resposta a infecções virais (Bernhardt *et al.*, 2024). Além disso, o gene *REG4* está relacionado à resposta inflamatória e pode interagir com processos inflamatórios induzidos pelos genes virais (Zheng; Xue; Zhang, 2022). O gene *WNK2*, já sugerido como um supressor tumoral pode estar positivamente regulado por *LMP-2A* e suas funções epigenéticas de ativação da *STAT3* já descritas nesse presente estudo (Hino *et al.*, 2009; Zhou *et al.*, 2019).

A expressão de *BRLF1*, que atua na fase lítica do ciclo viral, também pode influenciar na transcrição de mRNA (Long *et al.*, 2021). O gene viral *BZLF1*, também pertencente a fase lítica do vírus, não se encontra em nenhum dos clusters descritos acima, entretanto apresenta

correlação positiva com a maioria dos genes humanos DE. Essa correlação pode ser explicada pela função do *BZLF1*, já descrita no presente trabalho, onde o produto viral pode modificar a expressão de diversos genes do hospedeiro e pode suprimir as respostas do sistema inflamatório, facilitando assim a replicação do vírus (Long *et al.*, 2021).

Embora os genes humanos mencionados não tenham uma relação direta amplamente estudada com os genes virais, suas funções sugerem que eles podem ser modulados por proteínas virais, afetando a resposta do hospedeiro e a progressão da doença. Nesse sentido, são necessárias maiores pesquisas sobre essas correlações uma vez que interação entre os genes humanos e os genes virais do EBV é complexa e multifacetada.

7 CONCLUSÃO

Em suma, o presente estudo realizou a classificação molecular do EBV em adenocarcinoma gástrico por meio de dados de RNA-Seq, validados por testes imunohistoquímicos, e encontrou 8 amostras positivas. Além disso, ao analisar a expressão gênica dos pacientes encontramos 834 genes DE, sendo 93 deles com $AUC > 0.85$. Estes genes foram relacionados ao processo de fatores de crescimento, controle do ciclo celular, metástase e são associados a processos biológicos ligados a resposta imune e de defesa. Também foram analisadas as expressões dos genes virais do EBV nas amostras positivas, onde encontramos não só genes da fase de latência associada a EBVaCG, mas também há a expressão de genes da fase de replicação do vírus. Ao correlacionar essas expressões com os genes DE percebemos que há a modulação por meio dos genes virais que podem afetar a resposta do sistema imunológico do hospedeiro e a progressão da doença. Não obstante investigações adicionais ainda sejam necessárias para melhor compreensão da biologia do EBV no CG, o nosso estudo apresenta uma estratégia eficiente para a classificação molecular do subtipo EBVaCG baseada em NGS e elucida os efeitos do EBV na expressão gênica humana.

REFERÊNCIAS BIBLIOGRÁFICAS

ASHRAFIZADEH, M. et al. STAT3 Pathway in Gastric Cancer: Signaling, Therapeutic Targeting and Future Prospects. *Biology*, v. 9, n. 6, p. 126, 12 jun. 2020.

BAI, Y. et al. Efficacy and predictive biomarkers of immunotherapy in Epstein-Barr virus-associated gastric cancer. *Journal for ImmunoTherapy of Cancer*, v. 10, n. 3, p. e004080, 1 mar. 2022b.

BANERJEE, A. S.; PAL, A. D.; BANERJEE, S. Epstein–Barr virus-encoded small non-coding RNAs induce cancer cell chemoresistance and migration. *Virology*, v. 443, n. 2, p. 294–305, 1 set. 2013.

BARRA, W. F. et al. Gastric Cancer Microbiome. *Pathobiology*, v. 88, n. 2, p. 156–169, 2021.

BASS, A. J. et al. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*, v. 513, n. 7517, p. 202–209, set. 2014.

BERNHARDT, M. et al. Exploration of the Tumour Biological Significance of PCLO in Gastric Cancer: Results from a Large Central European Cohort. *Pathobiology*, v. 91, n. 3, p. 187–195, 7 nov. 2023.

BOROZAN, I. et al. Analysis of Epstein-Barr Virus Genomes and Expression Profiles in Gastric Adenocarcinoma. *Journal of Virology*, v. 92, n. 2, p. e01239-17, 15 jan. 2018.

CAMARGO, M. C. et al. Validation and calibration of next-generation sequencing to identify Epstein-Barr virus-positive gastric cancer in The Cancer Genome Atlas. *Gastric Cancer*, v. 19, n. 2, p. 676–681, abr. 2016.

CERVENKOVA, L. et al. Protein expression of ABCC2 and SLC22A3 associates with prognosis of pancreatic adenocarcinoma. *Scientific Reports*, v. 9, n. 1, p. 19782, 24 dez. 2019.

CHANG, X. et al. Bioinformatic Analysis Suggests That Three Hub Genes May Be a Vital Prognostic Biomarker in Pancreatic Ductal Adenocarcinoma. *Journal of Computational Biology*, v. 27, n. 11, p. 1595–1609, 1 nov. 2020.

CHENG, C.-W. et al. Sex hormone-binding globulin (SHBG) is a potential early diagnostic biomarker for gastric cancer. *Cancer Medicine*, v. 7, n. 1, p. 64–74, jan. 2018.

CHI, H.; PEPPER, M.; THOMAS, P. G. Principles and therapeutic applications of adaptive immunity. *Cell*, v. 187, n. 9, p. 2052–2078, 25 abr. 2024.

CORREA, P.; PIAZUELO, M. B. The gastric precancerous cascade: The gastric precancerous cascade. *Journal of Digestive Diseases*, v. 13, n. 1, p. 2–9, jan. 2012.

- DALEY, D. et al. Dectin-1 Activation on Macrophages by Galectin-9 Promotes Pancreatic Carcinoma and Peritumoral Immune-Tolerance. *Nature medicine*, v. 23, n. 5, p. 556–567, maio 2017
- DENG, W. et al. The novel secretory protein CGREF1 inhibits the activation of AP-1 transcriptional activity and cell proliferation. **The International Journal of Biochemistry & Cell Biology**, v. 65, p. 32–39, ago. 2015.
- DESHPANDE, D. et al. RNA-seq data science: From raw data to effective interpretation. **Frontiers in Genetics**, v. 14, 2023.
- DING, Y. et al. HAPLN3 inhibits apoptosis and promotes EMT of clear cell renal cell carcinoma via ERK and Bcl-2 signal pathways. **Journal of Cancer Research and Clinical Oncology**, v. 149, n. 1, p. 79–90, jan. 2023.
- DU, J. et al. LINC00858 knockdown inhibits gastric cancer cell growth and induces apoptosis through reducing WNK2 promoter methylation. **Cellular Oncology**, v. 43, n. 4, p. 709–723, 1 ago. 2020.
- EFTANG, L. L. et al. Up-regulation of CLDN1 in gastric cancer is correlated with reduced survival. **BMC Cancer**, v. 13, n. 1, p. 586, dez. 2013.
- EL-SHARKAWY, A.; AL ZAIDAN, L.; MALKI, A. Epstein–Barr Virus-Associated Malignancies: Roles of Viral Oncoproteins in Carcinogenesis. *Frontiers in Oncology*, v. 8, 2 ago. 2018.
- FANG, X. et al. Landscape of dietary factors associated with risk of gastric cancer: A systematic review and dose-response meta-analysis of prospective cohort studies. **European Journal of Cancer**, v. 51, n. 18, p. 2820–2832, dez. 2015.
- GANESAN, K. et al. Inhibition of Gastric Cancer Invasion and Metastasis by *PLA2G2A* , a Novel β -Catenin/TCF Target Gene. **Cancer Research**, v. 68, n. 11, p. 4277–4286, 1 jun. 2008.
- GOODACRE, N. et al. A Reference Viral Database (RVDB) To Enhance Bioinformatics Analysis of High-Throughput Sequencing for Novel Virus Detection. **mSphere**, v. 3, n. 2, p. e00069-18, 2018.
- GRADY, W. M.; YU, M.; MARKOWITZ, S. D. Epigenetic Alterations in the Gastrointestinal Tract: Current and Emerging Use for Biomarkers of Cancer. **Gastroenterology**, v. 160, n. 3, p. 690–709, fev. 2021.
- GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: A Cancer Journal for Clinicians**, v. 70, n. 4, p. 313–313, jul. 2020.
- GUANGCHUANG YU [AUT, C. clusterProfiler. [object Object], , 2017. Disponível em: <<https://bioconductor.org/packages/clusterProfiler>>. Acesso Goem: 2 ago. 2024

HAGN, M. et al. B cell-derived circulating granzyme B is a feature of acute infectious mononucleosis. **Clinical & Translational Immunology**, v. 4, n. 6, p. e38, jun. 2015.

HANAHAH, D.; WEINBERG, R. A. Hallmarks of Cancer: The Next Generation. **Cell**, v. 144, n. 5, p. 646–674, mar. 2011.

HIRABAYASHI, M. et al. Estimating the Global Burden of Epstein-Barr Virus–Associated Gastric Cancer: A Systematic Review and Meta-Analysis. **Clinical Gastroenterology and Hepatology**, v. 21, n. 4, p. 922–930.e21, abr. 2023.

HINO, R.; UOZAKI, H.; MURAKAMI, N.; USHIKU, T.; SHINOZAKI, A.; ISHIKAWA, S.; FUKAYAMA, M. Activation of DNA Methyltransferase 1 by EBV Latent Membrane Protein 2A Leads to Promoter Hypermethylation of PTEN Gene in Gastric Carcinoma. *Cancer Research*, v. 69, n. 7, p. 2766–2774, 2009. DOI: 10.1158/0008-5472.can-08-3070.

HONG, G. K. et al. Epstein-Barr Virus Lytic Infection Contributes to Lymphoproliferative Disease in a SCID Mouse Model. **Journal of Virology**, v. 79, n. 22, p. 13993–14003, 15 nov. 2005.

HUANG, S.-C. et al. Prognostic factors in Epstein-Barr virus-associated stage I–III gastric carcinoma: Implications for a unique type of carcinogenesis. **Oncology Reports**, v. 32, n. 2, p. 530–538, ago. 2014.

JANG, B.-G.; JUNG, E. J.; KIM, W. H. Expression of BamHI-A Rightward Transcripts in Epstein-Barr Virus-Associated Gastric Cancers. **Cancer Research and Treatment : Official Journal of Korean Cancer Association**, v. 43, n. 4, p. 250–254, dez. 2011.

JOHNSTON, F. M.; BECKMAN, M. Updates on Management of Gastric Cancer. **Current Oncology Reports**, v. 21, n. 8, p. 67, ago. 2019.

JORGOVANOVIC, D. et al. Roles of IFN- γ in tumor progression and regression: a review. **Biomarker Research**, v. 8, n. 1, p. 49, 29 set. 2020.

JU, M. K. et al. NSMF promotes the replication stress-induced DNA damage response for genome maintenance. **Nucleic Acids Research**, v. 49, n. 10, p. 5605–5622, 4 jun. 2021.

KANG, M.-S.; KIEFF, E. Epstein–Barr virus latent genes. **Experimental & Molecular Medicine**, v. 47, n. 1, p. e131–e131, 23 jan. 2015a.

KANG, M.-S.; KIEFF, E. Epstein–Barr virus latent genes. **Experimental & Molecular Medicine**, v. 47, n. 1, p. e131–e131, 23 jan. 2015b.

KONDO, A. et al. Loss of viral genome with altered immune microenvironment during tumour progression of EPSTEIN- BARR VIRUS - associated gastric carcinoma. **The Journal of Pathology**, v. 260, n. 2, p. 124–136, jun. 2023.

KURTH, I. et al. Therapeutic targeting of SLC6A8 creatine transporter suppresses colon cancer progression and modulates human creatine levels. **Science Advances**, v. 7, n. 41, p. eabi7511, 8 out. 2021.

LAUREN, P. THE TWO HISTOLOGICAL MAIN TYPES OF GASTRIC CARCINOMA: DIFFUSE AND SO-CALLED INTESTINAL-TYPE CARCINOMA. AN ATTEMPT AT A HISTO-CLINICAL CLASSIFICATION. **Acta Pathologica Et Microbiologica Scandinavica**, v. 64, p. 31–49, 1965.

LI, J. et al. Methylation and expression of Epstein–Barr virus latent membrane protein 1, 2A and 2B in EBV-associated gastric carcinomas and cell lines. **Digestive and Liver Disease**, v. 48, n. 6, p. 673–680, jun. 2016.

LI, J. et al. BST2 promotes gastric cancer metastasis under the regulation of HOXD9 and PABPC1. **Molecular Carcinogenesis**, v. 63, n. 4, p. 663–676, abr. 2024a.

LI, Q. et al. CTSC promoted the migration and invasion of glioma cells via activation of STAT3/SERPINA3 axis. **Gene**, v. 893, p. 147948, 30 jan. 2024b.

LIN, X. et al. Competitive Endogenous RNA Landscape in Epstein-Barr Virus Associated Nasopharyngeal Carcinoma. **Frontiers in Cell and Developmental Biology**, v. 9, 2021.

LIU, C. et al. A narrative review of the roles of muscle segment homeobox transcription factor family in cancer. **Annals of Translational Medicine**, v. 9, n. 9, p. 810–810, maio 2021.

LONG, X. et al. BRLF1-dependent viral and cellular transcriptomes and transcriptional regulation during EBV primary infection in B lymphoma cells. **Genomics**, v. 113, n. 4, p. 2591–2604, 1 jul. 2021.

LOVE, M. I.; HUBER, W.; ANDERS, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome Biology**, v. 15, n. 12, p. 550, 5 dez. 2014.

LU, Z. et al. Granzyme B Promotes Proliferation, Migration and EMT Process in Gastric Cancer. **Biochemical Genetics**, 27 maio 2024.

LUDDI, A. et al. Matrix metalloproteinases and their inhibitors in human cumulus and granulosa cells as biomarkers for oocyte quality estimation. **Fertility and Sterility**, v. 109, n. 5, p. 930- 939.e3, maio 2018.

MAO, Y. et al. High TNFSF13B expression as a predictor of poor prognosis in adrenocortical carcinoma. **Translational Andrology and Urology**, v. 10, n. 8, p. 3275–3285, ago. 2021.

MØLLER, M. et al. Heterogeneous patterns of DNA methylation-based field effects in histologically normal prostate tissue from cancer patients. **Scientific Reports**, v. 7, p. 40636, 13 jan. 2017.

MORALES-SANCHEZ, A.; FUENTES-PANANA, E. M. Epstein-Barr Virus-associated Gastric Cancer and Potential Mechanisms of Oncogenesis. **Current Cancer Drug Targets**, v. 17, n. 6, 30 jun. 2017a.

MORALES-SANCHEZ, A.; FUENTES-PANANA, E. M. Epstein-Barr Virus-associated Gastric Cancer and Potential Mechanisms of Oncogenesis. **Current Cancer Drug Targets**, v. 17, n. 6, 30 jun. 2017b.

MU, L. et al. Characterization of the Prognostic Values of CXCL Family in Epstein–Barr Virus Associated Gastric Cancer. **Oxidative Medicine and Cellular Longevity**, v. 2022, p. 1–24, 1 jun. 2022.

MURPHY, G. et al. Meta-analysis Shows That Prevalence of Epstein–Barr Virus-Positive Gastric Cancer Differs Based on Sex and Anatomic Location. **Gastroenterology**, v. 137, n. 3, p. 824–833, set. 2009.

NASEEM, M. et al. Outlooks on Epstein-Barr virus associated gastric cancer. **Cancer Treatment Reviews**, v. 66, p. 15–22, maio 2018.

OZCAN, S. C. et al. PFKFB2 regulates glycolysis and proliferation in pancreatic cancer cells. *Molecular and Cellular Biochemistry*, v. 470, n. 1–2, p. 115–129, jul. 2020.

PENG, K.; REN, X.; REN, Q. NcRNA-mediated upregulation of CAMK2N1 is associated with poor prognosis and tumor immune infiltration of gastric cancer. **Frontiers in Genetics**, v. 13, 25 ago. 2022.

POST, M. et al. The Transcription Factor ZNF683/HOBIT Regulates Human NK-Cell Development. **Frontiers in Immunology**, v. 8, p. 535, 2017.

ROBIN, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, v. 12, p. 77, 2011.

SAITO, M.; KONO, K. Landscape of EBV-positive gastric cancer. **Gastric Cancer**, v. 24, n. 5, p. 983–989, set. 2021.

SALMOND, R. J. Regulation of T Cell Activation and Metabolism by Transforming Growth Factor-Beta. **Biology**, v. 12, n. 2, p. 297, fev. 2023.

SALNIKOV, M. et al. Tumor-Infiltrating T Cells in EBV-Associated Gastric Carcinomas Exhibit High Levels of Multiple Markers of Activation, Effector Gene Expression, and Exhaustion. **Viruses**, v. 15, n. 1, p. 176, jan. 2023.

SHAKYA, M.; LO, C.-C.; CHAIN, P. S. G. Advances and Challenges in Metatranscriptomic Analysis. **Frontiers in Genetics**, v. 10, p. 904, 25 set. 2019.

SHENOY, A. R. et al. GBP5 Promotes NLRP3 Inflammasome Assembly and Immunity in Mammals. **Science**, v. 336, n. 6080, p. 481–485, 27 abr. 2012.

SHINOZAKI-USHIKU, A. et al. Profiling of Virus-Encoded MicroRNAs in Epstein-Barr Virus-Associated Gastric Carcinoma and Their Roles in Gastric Carcinogenesis. **Journal of Virology**, v. 89, n. 10, p. 5581–5591, 15 maio 2015.

SHINOZAKI-USHIKU, A.; KUNITA, A.; FUKAYAMA, M. Update on Epstein-Barr virus and gastric cancer (Review). **International Journal of Oncology**, v. 46, n. 4, p. 1421–1434, abr. 2015.

SHUTO, T. et al. Establishment of a Screening Method for Epstein-Barr Virus-Associated Gastric Carcinoma by Droplet Digital PCR. **Microorganisms**, v. 7, n. 12, p. 628, dez. 2019.

STRONG, M. J. et al. Differences in Gastric Carcinoma Microenvironment Stratify According to EBV Infection Intensity: Implications for Possible Immune Adjuvant Therapy. **PLoS Pathogens**, v. 9, n. 5, p. e1003341, 9 maio 2013a.

SUN, C.; MEZZADRA, R.; SCHUMACHER, T. N. Regulation and Function of the PD-L1 Checkpoint. *Immunity*, v. 48, n. 3, p. 434–452, mar. 2018.

SUN, K. et al. EBV-Positive Gastric Cancer: Current Knowledge and Future Perspectives. **Frontiers in Oncology**, v. 10, 2020.

SUNAKAWA, Y.; LENZ, H.-J. Molecular Classification of Gastric Adenocarcinoma: Translating New Insights from The Cancer Genome Atlas Research Network. **Current Treatment Options in Oncology**, v. 16, n. 4, p. 17, abr. 2015.

SUNG, H. et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. **CA: A Cancer Journal for Clinicians**, v. 71, n. 3, p. 209–249, maio 2021.

TAVAKOLI, A. et al. Association between Epstein-Barr virus infection and gastric cancer: a systematic review and meta-analysis. *BMC Cancer*, v. 20, p. 493, 1 jun. 2020

TANG, Y. et al. Expression of LMP and EBNA genes in Epstein-Barr virus-associated lymphomas in Hu-PBL/SCID mice. *Oncology Reports*, v. 35, n. 2, p. 905–911, 1 fev. 2016.

TIWARI, A. et al. Loss of HIF1A From Pancreatic Cancer Cells Increases Expression of PPP1R1B and Degradation of p53 to Promote Invasion and Metastasis. **Gastroenterology**, v. 159, n. 5, p. 1882- 1897.e5, nov. 2020.

TOKUNAGA, R. et al. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - A target for novel cancer therapy. *Cancer Treatment Reviews*, v. 63, p. 40–47, fev. 2018.

USUI, G. et al. DNA Methylation and Genetic Aberrations in Gastric Cancer. **Digestion**, v. 102, n. 1, p. 25–32, 2021

VIGANÒ, S. et al. Positive and Negative Regulation of Cellular Immune Responses in Physiologic Conditions and Diseases. **Clinical and Developmental Immunology**, v. 2012, p. 485781, 2012.

VISUS, C. et al. Identification of Hydroxysteroid (17 β) dehydrogenase type 12 (HSD17B12) as a CD8⁺ T-cell-defined human tumor antigen of human carcinomas. **Cancer Immunology, Immunotherapy : CII**, v. 60, n. 7, p. 919–929, 16 mar. 2011.

WANG, B. et al. Targeting regulatory T cells in gastric cancer: Pathogenesis, immunotherapy, and prognosis. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*, v. 158, p. 114180, fev. 2023.

WANG, J.-B. et al. Tumor Immunophenotyping-Derived Signature Identifies Prognosis and Neoadjuvant Immunotherapeutic Responsiveness in Gastric Cancer. **Advanced Science (Weinheim, Baden-Wurttemberg, Germany)**, v. 10, n. 15, p. e2207417, maio 2023.

WANG, X. et al. Expression of group IIA phospholipase A2 is an independent predictor of favorable outcome for patients with gastric cancer. **Human Pathology**, v. 44, n. 10, p. 2020–2027, out. 2013.

WANG, Z. et al. Identification of differential proteomics in Epstein-Barr virus-associated gastric cancer and related functional analysis. **Cancer Cell International**, v. 21, n. 1, p. 368, 12 jul. 2021.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, n. 1, p. 57–63, jan. 2009a.

WICKHAM H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.

WONG, Y. et al. Estimating the global burden of Epstein–Barr virus-related cancers. **Journal of Cancer Research and Clinical Oncology**, v. 148, n. 1, p. 31–46, jan. 2022b.

WOOD, Derrick E.; LU, Jennifer. Kraken2: ultrafast metagenomic sequence classification using exact alignments. 2021. Software. Disponível em: <<https://github.com/DerrickWood/kraken2>>. Acesso em: 02 ago. 2024.

WU, J. et al. Key genes associated with pancreatic cancer and their association with outcomes: A bioinformatics analysis. **Molecular Medicine Reports**, 3 jun. 2019.

XIANG, M. et al. A novel nomogram based on cell cycle-related genes for predicting overall survival in early-onset colorectal cancer. **BMC Cancer**, v. 23, n. 1, p. 595, 27 jun. 2023.

XIAO, Y. et al. Cathepsin C promotes breast cancer lung metastasis by modulating neutrophil infiltration and neutrophil extracellular trap formation. **Cancer Cell**, v. 39, n. 3, p. 423–437.e7, mar. 2021.

XU, J. et al. Matrix metalloproteinase expression and molecular interaction network analysis in gastric cancer. **Oncology Letters**, v. 12, n. 4, p. 2403–2408, out. 2016.

YAMASHITA, S. et al. Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2'-deoxycytidine treatment and oligonucleotide microarray. **Cancer Science**, v. 97, n. 1, p. 64–71, jan. 2006.

YANG, J. et al. Epstein–Barr virus-associated gastric cancer: A distinct subtype. **Cancer Letters**, v. 495, p. 191–199, dez. 2020a.

YANG, Y. et al. A comprehensive pan-cancer analysis on the immunological role and prognostic value of TYMP in human cancers. **Translational Cancer Research**, v. 11, n. 9, p. 3187–3208, set. 2022.

YASMIN, R. et al. Epigenetic Regulation of Inflammatory Cytokines and Associated Genes in Human Malignancies. **Mediators of Inflammation**, v. 2015, p. 1–8, 2015.

YIN, H. et al. Molecular mechanisms of EBV-driven cell cycle progression and oncogenesis. **Medical Microbiology and Immunology**, v. 208, n. 5, p. 573–583, out. 2019.

ZAVALA-VEGA, S. et al. Role of Epstein-Barr Virus in Glioblastoma. **Critical ReviewsTM in Oncogenesis**, v. 24, n. 4, p. 307–338, 2019.

ZEISEL, M. B.; DHAWAN, P.; BAUMERT, T. F. Tight junction proteins in gastrointestinal and liver disease. **Gut**, v. 68, n. 3, p. 547–561, mar. 2019.

ZHANG, C. et al. CXCL9/10/11, a regulator of PD-L1 expression in gastric cancer. **BMC Cancer**, v. 18, p. 462, 24 abr. 2018.

ZHANG, J. et al. LMP1 and EBNA2 constitute a minimal set of EBV genes for transformation of human B cells. **Frontiers in Immunology**, v. 14, p. 1331730, 19 dez. 2023.

ZHANG, Q.; XU, M. EBV-induced T-cell responses in EBV-specific and nonspecific cancers. **Frontiers in Immunology**, v. 14, 29 set. 2023.

ZHENG, H.-C.; XUE, H.; ZHANG, C.-Y. REG4 promotes the proliferation and anti-apoptosis of cancer. **Frontiers in Cell and Developmental Biology**, v. 10, p. 1012193, 2022

ZHOU, S.-L. et al. Genomic sequencing identifies WNK2 as a driver in hepatocellular carcinoma and a risk factor for early recurrence. **Journal of Hepatology**, v. 71, n. 6, p. 1152–1163, dez. 2019.

ZHOU, W. et al. Guanylate binding protein 4 shapes an inflamed tumor microenvironment and identifies immuno-hot tumors. **Journal of Cancer Research and Clinical Oncology**, v. 150, n. 2, p. 90, 12 fev. 2024.

ZUGUANG GU. ComplexHeatmap. [object Object], 2017. Disponível em: <<https://bioconductor.org/packages/ComplexHeatmap>>. Acesso em: 2 ago. 2024

ANEXO I



UNIVERSIDADE FEDERAL DO PARÁ NÚCLEO DE PESQUISAS EM ONCOLOGIA

Termo de Consentimento Livre e Esclarecido

Você está sendo convidado(a) para participar, como voluntário, em uma pesquisa. Após ser esclarecido(a) sobre as informações a seguir, no caso de aceitar fazer parte do estudo assine ao final deste documento, que está em duas vias e o mesmo será assinado e rubricado em todas as suas páginas. Uma delas é sua e a outra é do pesquisador responsável. Em caso de recusa você não será penalizado (a) de forma alguma. Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP) do Núcleo de Pesquisas em Oncologia, localizado no endereço: rua dos mundurucus, n. 4487, prédio da UNACON do Hospital Barros Barreto, telefone 3201-6778, email cep.npo@gmail.com.

Título de Projeto: Caracterização de potenciais biomarcadores e alvos terapêuticos do Adenocarcinoma Gástrico no estado do Pará.

O estudo é uma abordagem integrativa, que pretende, em uma amostra única, analisar concomitantemente, a expressão de transcritos humanos e não humanos, comparando-os e relacionando-os com os dados clínico-patológicos dos pacientes estudados. Com o intuito de caracterizar potenciais biomarcadores e alvos terapêuticos em adenocarcinomas gástricos de pacientes do Estado do Pará. Serão coletados pequenos fragmentos de tecido neoplásico removido por cirurgia ou biópsias (análise de pequenas amostras de tecido para o diagnóstico do tipo do tumor). Não havendo despesas pessoais para cada participante em qualquer fase do estudo. A obtenção das amostras para pesquisa não implicará riscos adicionais no tratamento ou na cirurgia.

Serão convidados a participar da pesquisa, pacientes com adenocarcinoma gástrico sem histórico de outro tipo de câncer. Não participarão do estudo, pacientes com história de outro tipo de câncer e doenças infecciosas. Estas coletas serão realizadas mediante sua livre aceitação e assinatura do deste Termo de Consentimento Livre e Esclarecido. Os possíveis riscos encontrados neste projeto estão relacionados a confidencialidade de dados do paciente para os estudos. Para diminuição deste risco, serão utilizadas estratégias para manter o sigilo de informações, como associar dados dos pacientes à códigos de identificação. Quanto ao procedimento de coleta, cirurgia e a biópsia gástrica são procedimentos indicados e realizados para estabelecimento de diagnóstico e/ou tratamento do adenocarcinoma gástrico, assim como, de outras doenças, sendo para tanto, necessária a remoção de material biológico relacionado à esta enfermidade. Parte do material retirado será encaminhada para as análises moleculares.

Não há benefício direto para o participante. Trata-se de estudo experimental investigando a expressão de transcritos humanos e não humanos. Os resultados dessa pesquisa buscam gerar informações que visam ampliar o conhecimento sobre a carcinogênese e o microbioma no câncer gástrico e posteriormente promover campos de pesquisa para o estabelecer biomarcadores de diagnósticos e prognósticos, e nortear o desenvolvimento de novas estratégias terapêuticas.

Os dados coletados serão mantidos sob sigilo, avaliados e utilizados apenas para esta pesquisa e você não será recompensado com qualquer tipo de pagamento por ter participado. Em qualquer

etapa do estudo você poderá ter acesso ao pesquisador responsável para esclarecimento de quaisquer dúvidas. O principal investigador é o Prof. Dr. Paulo Pimentel Assumpção, que pode ser encontrado no Núcleo de Pesquisas em Oncologia, 1º Piso da UNACON. Av. Mundurucus, 4487, Guamá, Belém-PA, CEP: 66073-000, telefone (91) 3201-6776, e-mail: assumpcaopp@gmail.com.

É garantida a liberdade de retirar o consentimento a qualquer momento e deixar de participar do estudo, sem qualquer prejuízo à continuidade do seu tratamento na instituição. As informações obtidas serão analisadas em conjunto com outros pacientes, não sendo divulgada a identificação de nenhum paciente. Esses materiais serão utilizados somente para este estudo. A eventual inclusão dos resultados em uma publicação científica será feita de modo a garantir o anonimato do participante.

Na eventualidade de qualquer dano pessoal causado direta ou indiretamente após procedimentos propostos neste estudo (anexo causal comprovado), o participante terá direito a tratamento médico na instituição, bem como indenizações estabelecidas.

Acredito ter sido suficientemente esclarecido a respeito das informações que li ou que foram lidas para mim, descrevendo um estudo “Caracterização de potenciais biomarcadores e alvos terapêuticos do Adenocarcinoma Gástrico no estado do Pará”. Eu discuti com

_____ sobre a minha decisão de participar do estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Concordo voluntariamente em participar deste estudo e poderei retirar meu consentimento a qualquer momento, antes ou durante o processo, sem penalidades, prejuízo ou perda de qualquer benefício que possa ter adquirido. Ressaltando que tal documento será emitido em duas vias, permanecendo uma com o participante e uma via com o pesquisador.

Data: _____/_____/_____

Nome do participante: _____

Assinatura do participante ou responsável legal: _____

Declaro que obtive de forma apropriada e voluntária o consentimento livre e esclarecido deste paciente ou representante legal para participação neste estudo.

Assinatura do responsável pelo estudo: _____

ANEXO II

Quadro 1 - Comparação das amostras submetidas aos dois testes. Amostras positivas possuíram leituras > 100.

Amostras	IHC	RNA-Seq
170111	Positivo	11435
170116	Negativo	0
170213	Negativo	0
170510	Negativo	0
170529	Negativo	8
170607	Negativo	0
180129	Negativo	0
180131	Negativo	0
180220	Negativo	0
180521	Negativo	0
180618	Negativo	2
180830	Positivo	58808
181101	Negativo	34
181210	Negativo	0
190121	Negativo	1
190128	Negativo	6
190415	Negativo	0
190417	Negativo	0
190506	Negativo	0
190611	Negativo	10
190624	Negativo	13
190701	Negativo	2
190708	Negativo	1
190724	Positivo	51919
190729	Positivo	7951
190731	Negativo	0
190807	Negativo	2
190829	Positivo	25456

191030	Negativo	0
191204	Negativo	0
200113	Negativo	2
200115	Positivo	1718
200122	Negativo	4
200316	Negativo	5
200325	Positivo	42172
200330	Negativo	0
200713	Negativo	0
200729	Negativo	-
200824	Negativo	-
210526	Positivo	404
210830	Negativo	-
210915	Negativo	-
211213	Negativo	0
210127	Negativo	-
210209	Negativo	-
210409	Negativo	0
210712	Negativo	0
211221	Negativo	2
220511	Negativo	0
220629	Negativo	-
220711	Negativo	0
220713	Positivo	4472
220822	Negativo	0
220829	Negativo	-
221228	Negativo	0
230125	Negativo	0
230607	Negativo	0
230612	Negativo	-
230614	Negativo	-

Fonte: Elaborada pelo autor.