UNIVERSIDADE DO ESTADO DO PARÁ
INSTITUTO EVANDRO CHAGAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA PARASITÁRIA NA AMAZÔNIA

ANDRÉ MILLER CORTÊZ LIMA

# ANÁLISE COMPARATIVA DA OCORRÊNCIA DE EPÍTOPOS LINEARES NA GLICOPROTEÍNA DE LISSAVÍRUS DOS FILOGRUPOS I E II

Belém/PA

2026

ANDRÉ MILLER CORTÊZ LIMA

# ANÁLISE COMPARATIVA DA OCORRÊNCIA DE EPÍTOPOS LINEARES NA GLICOPROTEÍNA DE LISSAVÍRUS DOS FILOGRUPOS I E II

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Biologia Parasitária na Amazônia da Universidade do Estado do Pará em parceria com o Instituto Evandro Chagas, como requisito para obtenção do título de mestre em Biologia Parasitária na Amazônia.

Orientador: Prof. Dr. Carlos Alberto Marques de Carvalho.

Coorientadora: Profa. Dra. Taciana Fernandes Souza Barbosa Coelho

Linha de pesquisa: Morfofisiologia e Genética dos Agentes Infecciosos, Parasitários e Vetores de Endemias na Amazônia.

Belém/PA

2026

ANDRÉ MILLER CORTÊZ LIMA

# ANÁLISE COMPARATIVA DA OCORRÊNCIA DE EPÍTOPOS LINEARES NA GLICOPROTEÍNA DE LISSAVÍRUS DOS FILOGRUPOS I E II

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Biologia Parasitária na Amazônia da Universidade do Estado do Pará em parceria com o Instituto Evandro Chagas, como requisito para obtenção de título de mestre em Biologia Parasitária na Amazônia, sob orientação do Prof. Dr. Carlos Alberto Marques de Carvalho e coorientação da Profa. Dra. Taciana Fernandes Souza Barbosa Coelho.

Linha de pesquisa: Morfofisiologia e Genética dos Agentes Infecciosos, Parasitários e Vetores de Endemias na Amazônia.

**BANCA EXAMINADORA**

Data da Aprovação: ___/__/____

_____

Prof. Dr. Carlos Alberto Marques de Carvalho

**Orientador** – Universidade do Estado do Pará (UEPA)

_____

Profa. Dra. Ana Cecília Ribeiro Cruz

**Membro Titular** – Universidade do Estado do Pará (UEPA)

_____

Profa. Dra. Livia Medeiros Neves Casseb

**Membro Titular** – Instituto Evandro Chagas (IEC)

_____

Prof. Dr. Samir Mansour Moraes Casseb

**Membro Titular** – Universidade Federal do Pará (UFPA)

_____

Prof. Dr. Leonardo Quintão Siravenha

**Membro Suplente** – Universidade do Estado do Pará (UEPA)

# SUMÁRIO

1 INTRODUÇÃO ..... 8
1.1 OS LISSAVÍRUS ..... 8
1.2 FILOGRUPO I ..... 8
1.2.1 Vírus da Raiva (*Lyssavirus rabies*) ..... 9
1.2.2 Lissavírus de Morcego Europeu 1 (*Lyssavirus hamburg*) ..... 10
1.2.3 Lissavírus de Morcego Europeu 2 (*Lyssavirus helsinki*) ..... 11
1.2.4 Vírus de Duvenhage (*Lyssavirus duvenhage*) ..... 12
1.2.5 Lissavírus de Morcego Australiano (*Lyssavirus australis*) ..... 13
1.3 FILOGRUPO II ..... 13
1.3.1 Lissavírus do Morcego de Lagos (*Lyssavirus lagos*) ..... 14
1.3.2 Vírus de Mokola (*Lyssavirus mokola*) ..... 15
1.4 CICLO REPLICATIVO DOS LISSAVÍRUS ..... 15
1.5 ASPECTOS IMUNOLÓGICOS DOS LISSAVÍRUS ..... 16
1.6 IMUNOGENICIDADE DA PROTEÍNA G DOS LISSAVÍRUS ..... 17
1.7 COMPLEXO PRINCIPAL DE HISTOCOMPATIBILIDADE ..... 19
1.7.1 HLA-I ..... 20
1.7.2 HLA-II ..... 21
1.7.3 HLA-III ..... 22
1.8 EPÍTOPOS DE CÉLULAS B E T ..... 22
1.9 BIOINFORMÁTICA NA PREDIÇÃO DE EPÍTOPOS LINFOCITÁRIOS ..... 23
2 OBJETIVOS ..... 26
2.1 GERAL ..... 26
2.2 ESPECÍFICOS ..... 26
3 MATERIAL E MÉTODOS ..... 27
3.1 ASPECTOS ÉTICOS E LEGAIS ..... 27
3.2 COLETA DAS SEQUÊNCIAS AMINOACÍDICAS ..... 27
3.3 DETERMINAÇÃO DA MATRIZ DE IDENTIDADE PERCENTUAL ..... 28
3.4 PREDIÇÃO DE EPÍTOPOS LINEARES DE CÉLULAS T $CD4^+$ E B ..... 28
4 RESULTADOS ..... 30
4.1 PERCENTUAL DE IDENTIDADE DAS SEQUÊNCIAS AMINOACÍDICAS ..... 30
4.2 PREDIÇÃO DE EPÍTOPOS LINEARES DE CÉLULAS T $CD4^+$ ..... 31
4.3 PREDIÇÃO DE EPÍTOPOS LINEARES DE CÉLULAS B ..... 34
5 DISCUSSÃO ..... 37

6 CONCLUSÃO ..... 40
7 BIBLIOGRAFIA ..... 42
8 ANEXO ..... 46

## RESUMO

Os lissavírus constituem um grupo de vírus neurotrópicos responsáveis por encefalite letal, com destaque para o vírus da raiva. A proteína G desses vírus desempenha papel central na infecção, sendo o principal alvo das imunidades adaptativas humoral e mediada por células. Nesse contexto, o presente estudo teve como objetivo analisar comparativamente a ocorrência de epítopos lineares de células B e T $CD4^+$ no ectodomínio da proteína G de lissavírus pertencentes aos filogrupos I (RABV, EBLV-1, EBLV-2, DUVV e ABLV) e II (LBV e MOKV), por meio de abordagens de bioinformática. As sequências aminoacídicas da proteína G foram obtidas por tradução conceitual a partir do genoma viral de referência disponível no GenBank, processadas no SignalP e no TMHMM para isolamento do ectodomínio, alinhadas no Clustal Omega para determinação de identidade percentual e submetidas ao EMBOSS Cons para geração de sequências consenso. A predição de epítopos de células T $CD4^+$ foi realizada com base na análise da afinidade de ligação a alelos de HLA-II pelo NetMHCIIpan com auxílio do Epitope-Evaluator, enquanto epítopos de células B foram identificados a partir da combinação de parâmetros de hidroficilidade, acessibilidade e flexibilidade no NPS@ PCProf, avaliando-se também o potencial mascaramento por N-glicanos no NetNGlyc. Os resultados evidenciaram a identidade aminoacídica de 76,71% a 83,79% no filogrupo I e 82,72% no filogrupo II. Observou-se maior densidade de epítopos para HLA-II no filogrupo I (0,22) em relação ao filogrupo II (0,18). Embora a antigenicidade média tenha sido maior em metades distintas do ectodomínio da proteína G em cada filogrupo de lissavírus, epítopos lineares de células B contidos em posições equivalentes dos filogrupos I e II (FKRKHF e TFKRKHFK, respectivamente) se sobrepuseram a peptídeos comumente ligados pelo alelo DRB1*15:01 de HLA-II (TYTNFVGYVTTTFKR, YTNFVGYVTTTFKRK e TNFVGYVTTTFKRKH) e não mascarados por N-glicanos, sugerindo relevância funcional em respostas de anticorpos dependentes de células T $CD4^+$. Em conjunto, esses achados contribuem para a compreensão de determinantes antigênicos da proteína G e fornecem subsídios para o desenvolvimento de estratégias vacinais e imunoprofiláticas mais abrangentes contra lissavírus.


**Palavras-chave:** Lissavírus; proteínas virais; epítopos; biologia computacional; imunidade.

## ABSTRACT

Lyssaviruses comprise a group of neurotropic viruses responsible for lethal encephalitis, most notably the rabies virus. The G protein of these viruses plays a central role in infection, being the main target of humoral and cell-mediated adaptive immunity. In this context, the present study aimed to comparatively analyze the occurrence of linear B and $CD4^+$ T cell epitopes in the ectodomain of the G protein of lyssaviruses belonging to phylogroups I (RABV, EBLV-1, EBLV-2, DUVV, and ABLV) and II (LBV and MOKV), using bioinformatics approaches. The amino acid sequences of the G protein were retrieved by conceptual translation from the reference viral genome available in GenBank, processed in SignalP and TMHMM for ectodomain isolation, aligned in Clustal Omega for percentage identity determination, and submitted to EMBOSS Cons for consensus sequence generation. Prediction of $CD4^+$ T cell epitopes was performed based on analysis of binding affinity to HLA-II alleles using NetMHCIIpan with the aid of Epitope-Evaluator, while B cell epitopes were identified from the combination of hydrophilicity, accessibility, and flexibility parameters in NPS@ PCProf, also evaluating the potential masking by N-glycans in NetNGlyc. The results showed amino acid identity of 76.71% to 83.79% in phylogroup I and 82.72% in phylogroup II. A higher density of HLA-II epitopes was observed in phylogroup I (0.22) compared to phylogroup II (0.18). Although the average antigenicity was higher in distinct halves of the G protein ectodomain in each lyssavirus phylogroup, linear B cell epitopes contained in equivalent positions of phylogroups I and II (FKRKHF and TFKRKHFK, respectively) overlapped to peptides commonly bound by the HLA-II DRB1*15:01 allele (TYTNFVGYVTTTFKR, YTNFVGYVTTTFKRK, and TNFVGYVTTTFKRKH) and not masked by N-glycans, suggesting functional relevance in $CD4^+$ T cell-dependent antibody responses. Taken together, these findings contribute to the understanding of antigenic determinants of the G protein and provide support for the development of more comprehensive vaccine and immunoprophylactic strategies against lyssaviruses.


**Keywords:** Lyssavirus; viral proteins; epitopes; computational biology; immunity.

# 1 INTRODUÇÃO

## 1.1 OS LISSAVÍRUS

O gênero *Lyssavirus* compreende um conjunto heterogêneo de vírus de RNA de fita simples e senso negativo (ssRNA$^-$) da família *Rhabdoviridae*, conhecidos por sua capacidade de causar encefalite potencialmente letal, uma condição característica da doença raiva. Sendo uma doença comumente associada à infecção por esses vírus, a raiva é considerada uma das enfermidades mais antigas documentadas pela humanidade, mencionada por Demócrito há aproximadamente 500 anos a.C. e por Aristóteles em 322 a.C., quando ambos observaram que cães saudáveis apresentavam comportamentos anormais após serem mordidos por cães doentes, além de sugerirem a possibilidade de transmissão da doença para os seres humanos (SCHNEIDER; SANTOS-BURGOA, 1994; REZENDE *et al*., 1997). Apesar do aumento significativo nas pesquisas científicas sobre essa temática nos últimos anos, ainda persistem várias lacunas no entendimento dos processos pelos quais os lissavírus invadem o organismo hospedeiro e das estratégias adaptativas que esses vírus utilizam para sobreviver dentro do corpo humano ou animal (BATISTA; FRANCO; ROEHE, 2007).

A disseminação dos lissavírus ocorre, predominantemente, por meio da mordida de animais que estejam infectados. Embora seja uma ocorrência rara, a transmissão também pode se dar por via aérea, mas essa forma de contágio não possui relevância epidemiológica significativa. Além disso, o contato direto com lesões abertas, secreções mucosas ou a realização de procedimentos médicos como transplantes de córnea e outros órgãos pode facilitar a propagação do vírus entre pessoas (BATISTA; FRANCO; ROEHE, 2007). O período de incubação da doença pode variar consideravelmente, sendo influenciado por fatores como a localização da mordedura (quanto mais próxima do sistema nervoso central a lesão, mais rapidamente o vírus se dissemina), a quantidade de vírus presente na saliva do animal infectado e o estado imunológico da pessoa que foi mordida ou ferida. Esses fatores podem alterar significativamente a velocidade e a intensidade da infecção (BATISTA; FRANCO; ROEHE, 2007).

## 1.2 FILOGRUPO I

Dentre os membros do filogrupo I, destacam-se o genótipo mundial 1 (vírus da raiva – RABV), os genótipos 5 (EBLV-1) e 6 (EBLV-2) do lissavírus de morcego europeu

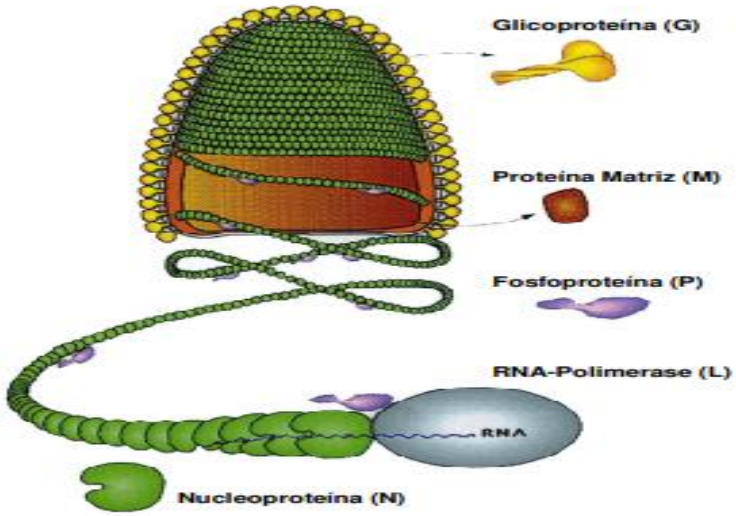
(EBLV), o genótipo africano 4 (vírus Duvenhage – DUVV) e o genótipo 7 do lissavírus de morcego australiano (ABLV) (EVANS *et al*., 2018).

### 1.2.1 Vírus da Raiva (*Lyssavirus rabies*)

O vírus da raiva (RABV, genótipo 1) é amplamente reconhecido como a espécie prototípica e a mais comum entre os membros do gênero *Lyssavirus*, sendo responsável por, aproximadamente, 59.000 mortes anuais. Esse vírus é altamente neurotrópico, o que significa que possui uma forte afinidade pelo sistema nervoso, e tem a capacidade de infectar praticamente todos os mamíferos terrestres (SCOTT; NEL, 2021).

A partícula do RABV possui uma forma característica de projétil de revólver, a qual se mantém preservada em todos os membros do gênero *Lyssavirus*, com uma largura aproximada de 75 nm, e um comprimento que pode variar entre 100 e 300 nm. Seu genoma de ssRNA$^-$ é empacotado dentro da partícula viral em uma simetria helicoidal, sendo envolvido por várias cópias da nucleoproteína (N) e associado à fosfoproteína (P) e à RNA-polimerase dependente de RNA (L), formando o nucleocapsídeo. Este nucleocapsídeo é, então, envolto pelas proteínas da matriz (M) e, finalmente, é cercado por um envelope de natureza lipídica derivado da membrana celular da célula hospedeira, no qual estão embutidas as glicoproteínas (G) (BATISTA; FRANCO; ROEHE, 2007) (Figura 1).

**Figura 1:** Morfologia do RABV.




**Fonte:** Adaptado de Dhulipala e Uversky (2022).

A proteína G, composta por 524 aminoácidos (para RABV, com discretas variações de tamanho entre os demais lissavírus) e apresentando um peso molecular entre

65 e 70 kDa, desempenha um papel crucial na interação do vírus com a célula hospedeira. Ela é responsável pela adsorção do vírus à célula e pela fusão do envelope viral com a membrana endossomal, um processo que ocorre após a internalização do vírus pela célula por meio de endocitose. A proteína G é a mais imunogênica das cinco proteínas codificadas pelo RNA viral, sendo a principal responsável pela indução de anticorpos neutralizantes. Isso ocorre devido à presença de um ectodomínio na estrutura tridimensional da glicoproteína, localizado na parte externa do envelope viral. Além disso, quando a proteína G é expressa em conjunto com as proteínas N e P, ela é capaz de estimular uma resposta imune adaptativa mediada por células. Essa resposta inclui o recrutamento de células T citotóxicas e, também, está envolvida no processo de brotamento de novos vírions, facilitando a disseminação do vírus dentro do organismo (BATISTA; FRANCO; ROEHE, 2007).

### 1.2.2 Lissavírus de Morcego Europeu 1 (*Lyssavirus hamburg*)

Na Europa, existem dois tipos de lissavírus de morcegos, conhecidos como lissavírus de morcegos europeus (EBLVs), que são os tipos 1 e 2 (correspondentes aos genótipos 5 e 6, respectivamente). Embora esses vírus infectem predominantemente morcegos, há registros raros de infecção em outros hospedeiros, como o gado doméstico, animais selvagens e, até mesmo, seres humanos. Em 1977, foram registradas quatro mortes causadas pela infecção por EBLV-1 (genótipo 5), sendo que as vítimas não haviam recebido nenhuma imunização profilática contra a raiva (FOOKS *et al*., 2003).

Além disso, a literatura científica aponta que os dados sobre a eficácia das vacinas atuais no combate à infecção por EBLV-1 são limitados, especialmente no que diz respeito à proteção imune cruzada entre os diferentes tipos de lissavírus. Nos primeiros estudos realizados, os isolados de EBLV-1 obtidos de morcegos foram analisados com o uso de anticorpos monoclonais, o que permitiu classificá-los no sorotipo 4, sendo identificados como um vírus com características semelhantes ao DUVV (KING; DAVIES; LAWRIE, 1990). Em seguida, um segundo grupo de vírus de morcegos foi identificado em espécies de morcegos como *Myotis dasycneme* e *Myotis daubentonii*.

Apesar de EBLV-1 e EBLV-2 compartilharem semelhanças genéticas e antigênicas, eles apresentam diferenças significativas entre si. Acredita-se que as diferentes linhagens de EBLV-1 tenham sido introduzidas em diversas regiões do norte da Europa de duas origens distintas. Especificamente, a linhagem EBLV-1a, que é a mais

recente, teria sido trazida do norte da África, chegando ao sul da Espanha antes de se espalhar para outras áreas. Os dois casos humanos com observações clínicas documentadas eram sugestivos da forma "furiosa" da raiva comum em pacientes infectados com o RABV clássico. Nos estágios iniciais da infecção, os sintomas clínicos típicos incluem dor persistente na região límbica, prurido no local da mordida e sensação de formigamento (parestesia). À medida que a infecção avança, uma fraqueza progressiva nos membros superiores é observada, que, posteriormente, estende-se para todos os membros, resultando em flacidez e hiporreflexia. A dor no pescoço é frequentemente relatada desde o início dos sintomas e tende a aumentar de intensidade, tornando-se contínua com o passar do tempo. Em alguns casos, pode ocorrer pneumotórax, o que requer ventilação assistida para o paciente. A progressão clínica é rápida, com o comprometimento neurológico intensificando-se à medida que o vírus se dissemina pelo sistema nervoso central. Outros sintomas clínicos incluem espasmos respiratórios, hiperexcitabilidade, ansiedade extrema, convulsões, salivação excessiva, distúrbios na percepção, confusão mental, e episódios de agressão, que, frequentemente, culminam em coma (ROINE *et al.,* 1988).

### 1.2.3 Lissavírus de Morcego Europeu 2 (*Lyssavirus helsinki*)

O EBLV-2 (genótipo 6) foi inicialmente identificado em 1985 por um biólogo suíço especializado no estudo de morcegos. Esse pesquisador estava envolvido em investigações científicas que se estendiam por diversos países, incluindo Finlândia, Suíça e Malásia, com foco na análise de diversas espécies de morcegos. Um aspecto importante do EBLV-2 é que ele apresenta um nível de homogeneidade de sequência genômica bastante elevado, o que significa que suas variações genéticas são bastante limitadas, além de possuir uma taxa evolutiva mais baixa quando comparado a outros tipos de lissavírus. Isso sugere que o EBLV-2 possui uma maior estabilidade genética, o que pode ter implicações no seu comportamento e na forma como ele se transmite (MCELHINNEY *et al*., 2018).

Em 1986, o EBLV-2 foi isolado na Suíça a partir de morcegos da espécie *M. daubentonii*, conhecidos popularmente como morcegos de Daubenton. Esta descoberta foi um marco importante, pois revelou a presença do vírus em uma nova população de morcegos, ampliando o conhecimento sobre sua distribuição e os possíveis vetores do patógeno. Entre 1987 e 2017, um total de 39 casos foram suspeitos de estarem relacionados ao EBLV-2, dos quais 34 foram efetivamente confirmados. Esse número,

embora significativo, é considerado relativamente baixo quando comparado ao número de casos associados ao EBLV-1, o que pode ser explicado por uma série de fatores, como, por exemplo, a vigilância epidemiológica ser menos eficiente em morcegos que não são sinantrópicos, como os morcegos de Daubenton, em comparação com morcegos de outras espécies, como aquelas do gênero *Eptesicus*. Além disso, outro fator que pode influenciar esses números é a baixa taxa de transmissão do EBLV-2 entre indivíduos da mesma espécie, como indicado por estudos experimentais realizados com morcegos de Daubenton. Esses estudos sugerem que, apesar do vírus estar presente em determinadas populações, ele pode não se espalhar com a mesma intensidade observada em outras espécies de lissavírus. Essas observações, descritas por pesquisadores como Johnson e colaboradores (2008), indicam que o EBLV-2 apresenta características específicas em termos de transmissão e evolução, as quais são importantes para a compreensão dos riscos que ele pode representar para a saúde pública e para o monitoramento das populações de morcegos.

### 1.2.4 Vírus de Duvenhage (*Lyssavirus duvenhage*)

A infecção pelo DUVV (genótipo 4) é uma causa conhecida de encefalite letal, afetando tanto humanos quanto animais. Embora o DUVV seja amplamente prevalente entre as populações de morcegos na África, os casos de infecção humana são extremamente raros. Até o momento, foram documentados apenas três casos fatais de infecção em humanos: dois na África do Sul e um no Quênia. Esses episódios são considerados extremamente incomuns, refletindo a baixa incidência de transmissão do vírus para seres humanos, apesar de sua disseminação significativa entre morcegos na região. Embora o DUVV seja geneticamente muito semelhante ao RABV, ele provoca lesões distintas em seres humanos (PAWESKA *et al.*, 2006).

A infecção pelo RABV é tipicamente associada à presença de corpos de inclusão citoplasmáticos eosinofílicos, conhecidos como corpos de Negri, em neurônios, e a inflamação não costuma ser um dos principais aspectos da doença. Por outro lado, nas infecções humanas causadas pelo DUVV, não foram observados corpos de Negri e a inflamação foi consideravelmente mais proeminente, acompanhada de morte celular, que se destaca como uma característica marcante dessa infecção (VAN THIEL *et al.*, 2009). Esses resultados reforçam as diferenças patológicas entre os dois vírus, mesmo com sua proximidade genética.

### 1.2.5 Lissavírus de Morcego Australiano (*Lyssavirus australis*)

A infecção humana pelo ABLV (genótipo 7) é considerada extremamente rara, porém, é uma condição letal, sem que haja, até o momento, um tratamento comprovadamente eficaz para a cura. O diagnóstico dessa infecção, assim como para outros tipos de lissavírus, é notoriamente desafiador e, frequentemente, realizado de forma tardia. Isso pode ser observado nos exames neurológicos de imagem, nos testes sorológicos e na análise do líquido cefalorraquidiano (LCR), os quais, em muitos casos, não apresentam alterações significativas nas fases iniciais da doença rábica, dificultando uma identificação precoce da patologia. A reação em cadeia da polimerase com transcrição reversa (RT-PCR) tem se mostrado uma ferramenta valiosa e mais sensível para a identificação do ABLV. Para a realização do diagnóstico, é recomendada a coleta de uma biópsia da pele do paciente, pois essa amostra apresenta maior sensibilidade do que as coletas de saliva ou de LCR, que tendem a ser menos confiáveis para a detecção precoce do vírus (FRANCIS *et al*., 2014).

A transmissão do ABLV, de forma similar à de outras espécies de lissavírus, ocorre, principalmente, por meio do contato direto com a saliva de um animal infectado. Esse contato pode ocorrer através de mordidas, arranhões ou até mesmo pela exposição a feridas abertas, embora o risco de transmissão seja mais elevado com a mordida de um animal contaminado. Embora a endemicidade do ABLV na Austrália não seja tão alta quando comparada a de outros lissavírus, as infecções em humanos ainda apresentam uma taxa de letalidade extremamente alta, o que torna a doença particularmente grave e difícil de tratar. A implacabilidade da letalidade dessa infecção continua sendo uma preocupação, especialmente diante da ausência de terapias eficazes, e reforça a importância de vigilância contínua e medidas preventivas para minimizar os riscos de exposição ao vírus (FRANCIS *et al*., 2014).

## 1.3 FILOGRUPO II

Dentre os membros do filogrupo II, destacam-se o genótipo 2 (lissavírus do morcego de Lagos – LBV) e o genótipo 3 (vírus Mokola – MOKV) (HASSAN *et al.*, 2001).

### 1.3.1 Lissavírus do Morcego de Lagos (*Lyssavirus lagos*)

Até a década de 1950, acreditava-se que o RABV era o único agente responsável pela doença conhecida como raiva. No entanto, essa percepção mudou significativamente em 1956, quando foi identificado, pela primeira vez, outro vírus relacionado à raiva, o LBV (genótipo 2) (BOULGER; PORTERFIELD, 1958). A descoberta ocorreu na Nigéria e representou um marco importante para a compreensão da complexidade da doença, uma vez que revelou que outros vírus, além do RABV, também poderiam ser responsáveis por manifestações semelhantes à raiva em diferentes espécies de animais, especialmente os morcegos. Essa descoberta ampliou o entendimento sobre os agentes causadores da doença e abriu novas portas para a pesquisa e o estudo de outros vírus pertencentes ao mesmo grupo, que estão envolvidos em infecções em seres humanos e animais.

A vasta diversidade de lissavírus observada na África levou à formulação da hipótese de que esses vírus podem ter tido sua origem no continente africano, onde evoluíram ao longo de milhares de anos em estreita interação com seus hospedeiros primários, os morcegos. Acredita-se que, ao longo desse extenso período, os lissavírus e os morcegos tenham coevoluído, desenvolvendo uma relação complexa e adaptativa. Essa teoria sugere que, à medida que os morcegos se espalharam por diversas regiões, os lissavírus também se diversificaram, tornando-se endêmicos em várias espécies de morcegos, o que pode ter facilitado a transmissão e a persistência dos vírus ao longo do tempo. Esse processo de coevolução teria sido fundamental para a manutenção e disseminação dos lissavírus, contribuindo para a sua diversidade genética e a adaptação a diferentes espécies de morcegos em várias partes do mundo. O primeiro genoma completo do LBV foi publicado somente em 2008, proporcionando uma melhor compreensão sobre a estabilidade genética do vírus, que parece manter uma consistência ao longo do tempo e do espaço (COERTSE *et al.,* 2021).

Um aspecto particularmente interessante dos vírus pertencentes à linhagem A do LBV, que é a única linhagem para a qual múltiplos genomas completos estão disponíveis, é a observação de uma altíssima estabilidade genética. Foram detectados vírus dessa linhagem em duas regiões distintas da África – África Ocidental e África Oriental – com uma diferença de 28 anos entre as amostras. Apesar dessa diferença temporal considerável, a divergência de nucleotídeos entre os genomas sequenciados foi inferior a 1,5% (COERTSE *et al.,* 2021). Esse baixo índice mutacional indica que o LBV pode ter uma taxa de evolução relativamente lenta, o que sugere uma adaptação muito eficaz ao

seu hospedeiro e ao ambiente, mantendo a sua identidade genética muito bem preservada ao longo de décadas e de diferentes locais geográficos.

### 1.3.2 Vírus de Mokola (*Lyssavirus mokola*)

O MOKV (genótipo 3), relacionado à raiva, é uma das espécies reconhecidas dentro do gênero *Lyssavirus* capazes de causar encefalite fatal, conforme destacado por Walker e colaboradores (2018). Em 1963, Jack Schmidt, um pesquisador da Unidade de Pesquisa Médica Naval no Cairo, conseguiu isolar um vírus a partir de dois mosquitos da espécie *Mansonia uniformis*, que foram capturados no Sudão. Esse vírus recebeu o nome de *Obodhiang*, em referência ao local exato onde foi encontrado. Schmidt enviou amostras do vírus para a Universidade de Yale, esperando que o agente fosse analisado e corretamente identificado. Em 1967, Vernon Lee isolou um outro vírus, denominado *Kotonkan*, e, no ano seguinte, em 1968, foi a vez de Graham Kemp isolar o MOKV. Essas descobertas foram cruciais para o avanço do entendimento sobre os vírus relacionados à raiva, especialmente no que se refere à diversidade dos lissavírus, ampliando o conhecimento sobre os patógenos responsáveis por doenças em humanos e animais, principalmente na África. O hospedeiro reservatório do MOKV ainda é desconhecido; desde a sua descoberta em 1960, foi sugerido que o vírus tem uma predileção por pequenos mamíferos terrestres, uma hipótese que se baseia nos isolamentos de MOKV encontrados em musaranhos (*Crocidura* sp.) e em um roedor, o *Lophuromys sikapusi* (MCMAHON *et al.,* 2021). Essa teoria sustenta a ideia de que o MOKV pode se propagar, principalmente, entre mamíferos não voadores, embora o reservatório definitivo ainda não tenha sido identificado.

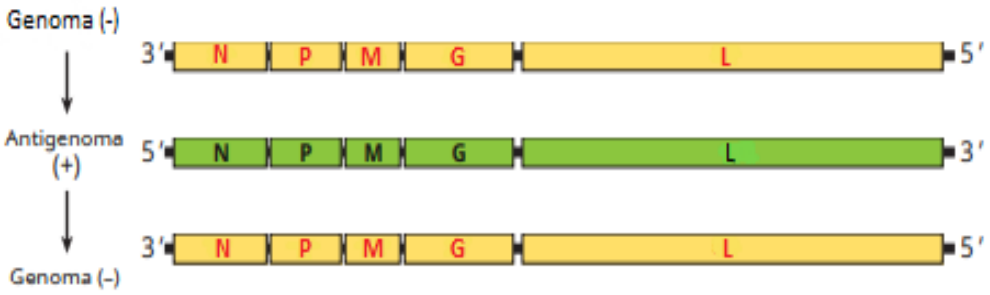
## 1.4 CICLO REPLICATIVO DOS LISSAVÍRUS

Como primeiro passo do ciclo replicativo, a proteína G, em sua forma trimérica, interage com receptores específicos, como o receptor nicotínico de acetilcolina (nAChR), a molécula de adesão celular neuronal (NCAM) e o receptor de neurotrofinas p75 (p75NTR), presentes na célula hospedeira (LIAN; HUEFFER; WELTZIN, 2022). Essa interação promove a internalização do vírion por meio de endocitose, seguida pela fusão do envelope viral com a membrana endossomal. Esse processo de fusão é mediado por um domínio específico localizado entre os resíduos 103 e 179 da forma madura da proteína G (DURRER *et al.*, 1995).

O genoma viral, de fita simples e senso negativo, é então liberado no citosol, onde é transcrito pela proteína L em cinco RNAs mensageiros (mRNAs). Esses mRNAs, capeados e poliadenilados, codificam as proteínas virais N, P, M, G e L. A proteína G, no entanto, sofre ainda modificações pós-traducionais, sendo glicosilada no complexo de Golgi e, posteriormente, transportada até a membrana plasmática. A replicação do RNA viral só tem início após a tradução dos mRNAs, sendo um processo regulado pelas concentrações citoplasmáticas de proteína N e RNA. Assim, quando os níveis de proteína N aumentam significativamente, ocorre a interrupção da transcrição e tradução, sinalizando a transição para a etapa de replicação do RNA (BATISTA; FRANCO; ROEHE, 2007).

Inicialmente, o genoma viral é replicado na forma de RNA de senso positivo (anti-genoma), que serve como molde para a síntese de novas fitas de RNA de senso negativo, destinadas à formação dos novos vírions (Figura 2). Durante a montagem viral, as proteínas N, P e L participam da encapsidação dos genomas recém-sintetizados. A proteína M envolve o complexo ribonucleoproteico (RNP), que é direcionado para regiões da membrana plasmática ricas em proteína G. Nessa etapa, a proteína M promove o enovelamento do RNP, conferindo-lhe uma estrutura helicoidal semelhante a uma mola, e inicia-se, então, a formação do envelope viral. Ao término do processo, os novos vírions são liberados por brotamento, sem causar lise das células infectadas (BATISTA; FRANCO; ROEHE, 2007).

**Figura 2:** Replicação do genoma dos lissavírus.




**Fonte:** DAVIS; RALL; SCHNELL, 2015.

## 1.5 ASPECTOS IMUNOLÓGICOS DOS LISSAVÍRUS

A resposta imune protetora contra lissavírus é predominantemente mediada por anticorpos neutralizantes direcionados contra a proteína G viral, principal determinante antigênico da partícula viral (FABER *et al.*, 2009). A proteína G é responsável pela

ligação aos receptores celulares e pela fusão de membranas, sendo também o principal alvo de anticorpos neutralizantes induzidos por vacinação. Estudos recentes reforçam que a indução de títulos adequados de anticorpos neutralizantes anti-G é o principal correlato de proteção contra a infecção (RUPPRECHT; HANLON; HEMACHUDHA, 2002; BANYARD *et al.*, 2013).

Entretanto, diferenças antigênicas entre filogrupos impactam a eficácia da proteção cruzada. Lissavírus do filogrupo II, como LBV e MOKV, apresentam menor neutralização por anticorpos induzidos por vacinas baseadas em RABV. Essa divergência antigênica está associada a variações estruturais na proteína G, que alteram epítopos reconhecidos por anticorpos neutralizantes (BANYARD *et al.*, 2013). Sabe-se, também, que, do ponto de vista imunológico, a infecção por lissavírus apresenta particularidades relevantes: a replicação inicial em tecido muscular é seguida por invasão do sistema nervoso periférico e posterior disseminação ao sistema nervoso central, onde a resposta inflamatória é limitada. A evasão imunológica é facilitada tanto pela natureza neurotrópica do vírus quanto por mecanismos de modulação da resposta mediada por interferon (RUPPRECHT; HANLON; HEMACHUDHA, 2002).

## 1.6 IMUNOGENICIDADE DA PROTEÍNA G DOS LISSAVÍRUS

A proteína G, sendo a mais imunogênica das cinco proteínas codificadas pelo ssRNA$^-$ dos lissavírus, desempenha um papel crucial na indução de anticorpos neutralizantes. Essa característica se deve, principalmente, à presença de um ectodomínio em sua estrutura tridimensional, que está localizado na parte externa do envelope viral, facilitando sua interação com o sistema imune. Além disso, quando a proteína G é expressa em conjunto com as proteínas N e P, ela tem a capacidade de ativar uma resposta imune adaptativa mediada por células. Esse processo resulta no recrutamento de células T citotóxicas, responsáveis pela destruição de células infectadas. Badrane e Tordo (2001) também constataram uma diferença na patogenicidade entre os membros do filogrupos I e II. O estudo demonstrou que o primeiro grupo apresenta competência para infecções nas vias intracerebrais e intramusculares, enquanto o segundo apresenta proficiência infecciosa apenas na via intracerebral, o que pode estar associado a uma mutação no resíduo R333 da proteína G do filogrupo I, que apresenta uma substituição R333D no filogrupo II.

Sabe-se, também, que sítios mais antigênicos presentes na estrutura da referida proteína se encontram na região do seu ectodomínio, conforme foi demonstrado no estudo de Evans e colaboradores (2018). As proteínas G presentes nos lissavírus possuem uma semelhança de 57% a 78% entre diferentes espécies, o que exerce um impacto considerável na proteção cruzada, especialmente após a administração da vacina contra o RABV (CALLAWAY *et al.*, 2022). Além disso, um estudo conduzido por Shi e colaboradores (2022) indicou que a vacina desenvolvida especificamente para o controle do RABV é capaz de fornecer uma proteção eficaz contra esse vírus quando o título de anticorpos neutralizantes alcança um valor superior a 0,5 unidade internacional (UI).

Sabe-se que a vacinação antirrábica constitui uma das estratégias mais eficazes já estabelecidas para a prevenção de doença rábica humana. Contudo, as formulações atuais dos imunizantes contra o RABV não conseguem induzir a formação de anticorpos neutralizantes suficientes nem oferecem uma proteção eficaz contra os lissavírus pertencentes ao filogrupo II, o que representa uma lacuna significativa nas estratégias de prevenção para essas infecções virais (BEN HAMED *et al.*, 2024). Estudos também demonstram que a combinação entre vacinação e imunização passiva com imunoglobulinas ou anticorpos monoclonais amplia a proteção em exposições de maior risco (NAGARAJAN, MARISSEN e RUPPRECHT, 2014).

A base da proteção imunológica está no reconhecimento de epítopos antigênicos, principalmente da glicoproteína G viral, capazes de estimular a produção de anticorpos neutralizantes. No entanto, embora historicamente o termo “vacina antirrábica” esteja associado à proteção contra a infecção pelo vírus rábico clássico, o gênero *Lyssavirus* compreende um conjunto mais amplo de vírus geneticamente e antigenicamente distintos. Nesse contexto, a eficácia protetora das vacinas atuais depende do grau de similaridade antigênica entre o vírus vacinal e o vírus circulante. De modo geral, a proteção cruzada conferida pelas vacinas atuais mostra-se mais consistente contra lissavírus filogeneticamente próximos do RABV, particularmente aqueles pertencentes ao filogrupo I (BROOKES *et al.,* 2005; EVANS *et al.,* 2012).

O estudo de Echevarría e colaboradores (2019) demonstrou que as vacinas antirrábicas atualmente disponíveis não conferem imunidade protetora adequada contra lissavírus pertencentes ao filogrupos II. Entre esses vírus destacam-se o *Lyssavirus lagos* (LBV), *Lyssavirus mokola* (MOKV), cujas diferenças antigênicas reduzem substancialmente o reconhecimento por anticorpos induzidos pelas formulações vacinais

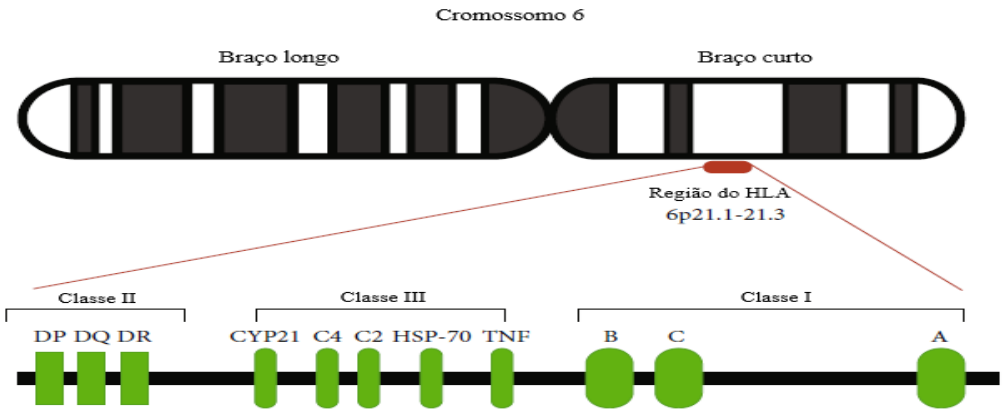
baseadas em RABV. Esses achados reforçam que a neutralização sorológica e a proteção cruzada tendem a ser limitadas quando há maior distância evolutiva entre os lissavírus (ECHEVARRÍA *et al.,* 2019).

Esse cenário possui implicações importantes para a vigilância epidemiológica, para a saúde pública e para o desenvolvimento de novas estratégias vacinais. Embora as vacinas atuais permaneçam altamente eficazes contra a raiva clássica e contra diversos membros do filogrupo I, a circulação de lissavírus antigenicamente divergentes evidencia uma lacuna de proteção imunológica. Estudos experimentais e revisões recentes têm destacado, portanto, a necessidade de formulações vacinais de espectro ampliado, capazes de induzir respostas neutralizantes mais abrangentes e potencialmente eficazes contra múltiplos filogrupos de lissavírus. Assim, a compreensão das relações antigênicas e filogenéticas dentro do gênero *Lyssavirus* torna-se essencial para orientar futuras abordagens de prevenção, vigilância e desenvolvimento de vacinas pan-lyssavirus (EVANS *et al*., 2012; SERVAT *et al*., 2019).

## 1.7 COMPLEXO PRINCIPAL DE HISTOCOMPATIBILIDADE

O complexo principal de histocompatibilidade (MHC) se trata de uma região altamente polimórfica do genoma de vertebrados, responsável por codificar um conjunto de glicoproteínas localizadas na superfície celular, proficientes no reconhecimento de estruturas antigênicas de natureza endógena ou exógena. Essas glicoproteínas desempenham um papel fundamental na apresentação antígenos às células T, visando o acionamento de seus mecanismos efetores. Em seres humanos, denominado antígeno leucocitário humano (HLA), esta molécula se apresenta em três configurações: HLA-I, HLA-II e HLA-III, e encontra-se estrategicamente alocada entre as regiões centromérica e telomérica do cromossomo 6 (Figura 3). Desse modo, fica evidente que o MHC atua protagonizando a regulação da resposta imune adaptativa mediada por células, apesar dos mecanismos moleculares desta interação ainda não estarem robustamente consolidados na literatura. O elevado polimorfismo no *locus* do HLA contribui para a diversidade genética da espécie e determina a suscetibilidade ou proteção dos indivíduos a diferentes processos infecciosos, conforme observado na tuberculose, hanseníase e em infecções de etiologia viral (SILVA-PINTO; SANTOS, 2011).

**Figura 3:** *Locus* cromossômico do HLA.

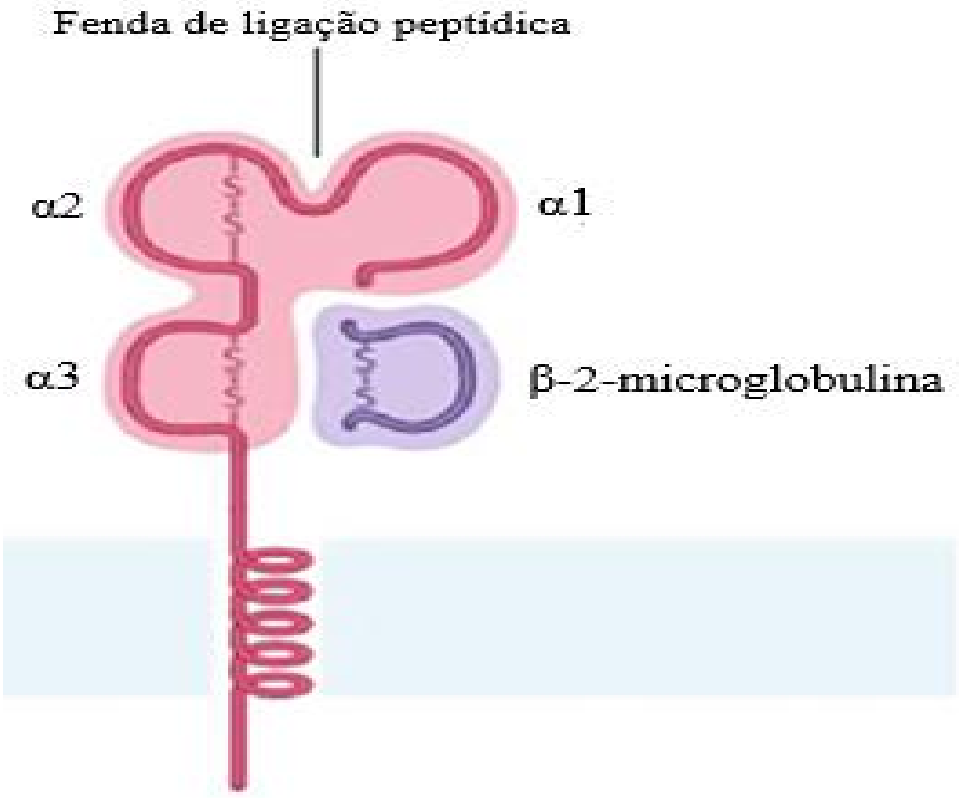



**Fonte:** Adaptado de Medhasi e Chantratita, 2022.

### 1.7.1 HLA-I

O HLA-I, que tem como principal função participar da resposta imune adaptativa contra patógenos que se encontram no interior das células, é uma glicoproteína transmembrana presente na superfície de praticamente todas as células nucleadas. A estrutura molecular dessa classe de HLA é composta por uma cadeia polipeptídica heterodimérica, o que significa que ela é formada por duas unidades monoméricas distintas que estão unidas por ligações químicas. Essas duas cadeias são: uma cadeia α pesada e uma cadeia β-2-microglobulina, sendo que a cadeia α é maior e desempenha um papel central na estrutura funcional da molécula. A cadeia α possui três domínios extracelulares, denominados α1, α2 e α3, além de uma porção transmembrana e uma região C-terminal que fica no citoplasma da célula. Os domínios α1 e α2, em conjunto, se dobram de uma maneira específica, formando um sulco ou fenda na estrutura da molécula, conhecida como "fenda de ligação peptídica" (Figura 4). Essa fenda é fundamental, pois é nela que ocorre o reconhecimento e a ligação dos peptídeos antigênicos, ou seja, as pequenas sequências de proteínas derivadas de patógenos que serão apresentadas às células T citotóxicas. Por outro lado, a cadeia β-2-microglobulina, que é menor em relação à cadeia α, está associada principalmente à porção α3 da cadeia pesada e desempenha um papel crucial na manutenção da estabilidade da molécula. A interação entre essas duas cadeias é vital para o funcionamento adequado do HLA-I, garantindo que a molécula tenha a conformação necessária para desempenhar suas funções de forma eficaz (MEDHASI; CHANTRATITA, 2022).

**Figura 4:** Estrutura química do HLA-I.

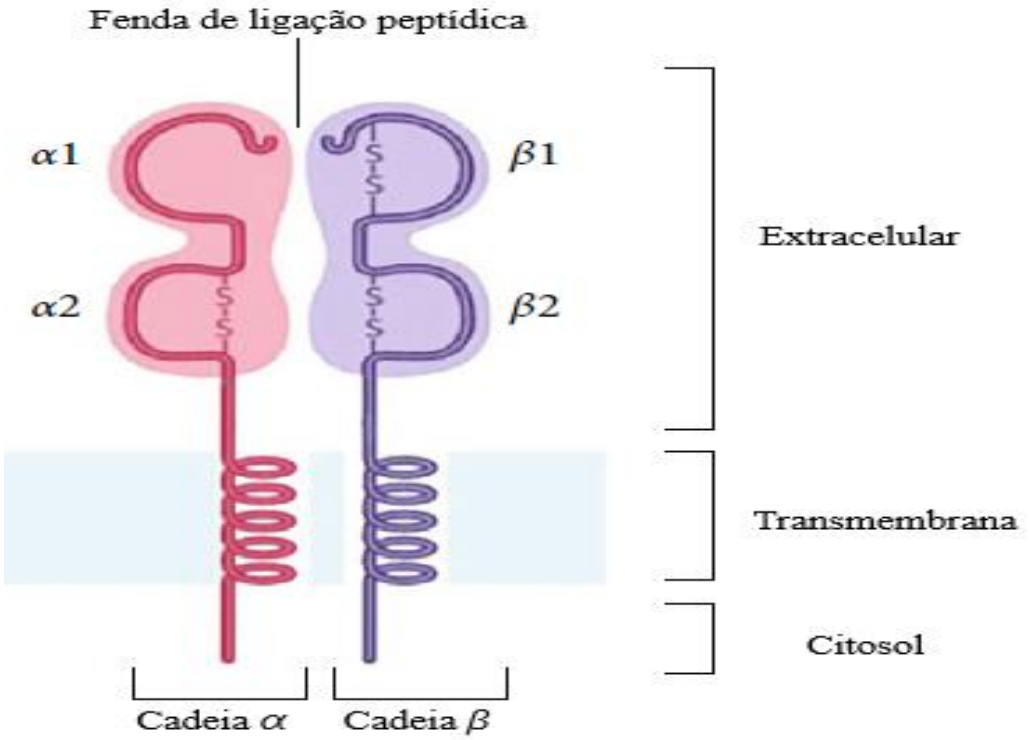



**Fonte:** Adaptado de Medhasi e Chantratita, 2022.

### 1.7.2 HLA-II

O HLA-II é uma molécula que se encontra na superfície de células apresentadoras de antígenos profissionais, como os macrófagos, as células dendríticas e as células B. Essas células desempenham um papel essencial na resposta imunológica, pois são responsáveis por processar e apresentar fragmentos de antígenos às células T auxiliares, ativando assim a resposta imune adaptativa mediada por células. As moléculas de HLA-II possuem uma estrutura composta por duas cadeias polipeptídicas, denominadas cadeia α e cadeia β. Cada uma dessas cadeias se dobra em dois domínios distintos: para a cadeia α, os domínios são denominados α1 e α2, e para a cadeia β, os domínios são chamados de β1 e β2 (Figura 5). Vale ressaltar que os domínios α2 e β2 são particularmente importantes, pois são altamente conservados em termos evolutivos, o que significa que sua estrutura é mantida de maneira muito similar entre diferentes indivíduos e espécies. Além disso, esses domínios desempenham um papel crucial na interação com os receptores das células T auxiliares, facilitando a comunicação entre as células apresentadoras de antígeno e tais células, o que é fundamental para a ativação do sistema imune (MEDHASI; CHANTRATITA, 2022).

**Figura 5:** Estrutura química do HLA-II.




**Fonte:** Adaptado de Medhasi e Chantratita, 2022.

### 1.7.3 HLA-III

A terceira configuração molecular do HLA, conhecida como HLA-III, corresponde a uma área específica que desempenha um papel crucial na codificação de diversos componentes envolvidos na resposta inflamatória, bem como proteínas associadas ao estresse térmico e elementos integrantes do sistema complemento. Esta organização molecular é fundamental para o funcionamento do sistema imunológico, particularmente em processos que envolvem inflamação e defesa contra infecções (MEDHASI; CHANTRATITA, 2022).

## 1.8 EPÍTOPOS DE CÉLULAS B E T

A avaliação experimental da eficácia de epítopos candidatos é um processo oneroso, demorado e tecnicamente complexo. Nesse contexto, o uso de abordagens computacionais para o *design* racional de vacinas, baseadas em ferramentas e bancos de dados de bioinformática, tem despertado crescente interesse. Essas ferramentas permitem a análise sistemática de sequências proteicas, o mapeamento e a predição *in silico* de epítopos conservados e imunogênicos de células B e T, além da avaliação de sua afinidade com moléculas do MHC. Assim, a bioinformática desempenha um papel central na

seleção de epítopos promissores e no desenvolvimento de estratégias vacinais mais eficientes e abrangentes, como a proposta de uma nova vacina pan-*Lyssavirus*, reduzindo custos, tempo e etapas experimentais iniciais (ÁVILA, 2011).

Epítopos são caracterizados como conjuntos de aminoácidos provenientes de um antígeno proteico que se ligam aos receptores das células B (imunoglobulinas) ou aos receptores das células T, promovendo a ativação de uma resposta imunológica. Sabe-se, também, que o reconhecimento antigênico executado por estas células ocorre de forma diferente. O receptor das células B (BCR) é capaz de se ligar diretamente a epítopos livres, enquanto o receptor das células T (TCR) reconhece antígenos previamente processados pela ação de um proteassoma ou fagolisossomo e associados a uma molécula de MHC (ÁVILA, 2011).

Os epítopos podem ser classificados entre contínuos ou descontínuos. O epítopo contínuo, também conhecido como sequencial ou linear, corresponde a um trecho contínuo da sequência de aminoácidos da proteína, ou seja, com aminoácidos organizados um após o outro. Por outro lado, o epítopo descontínuo, também chamado de conformacional, leva em consideração a topografia da proteína e não uma sequência de resíduos lineares. É formado por diferentes fragmentos distribuídos ao longo da sequência proteica, e que se aproximam espacialmente durante o processo de enovelamento da proteína (MENDEL *et al.*, 1996)

A identificação de epítopos é fundamental para a compreensão do reconhecimento molecular da proteína em estudo, principalmente ao considerar a sua capacidade de desencadear mecanismos efetores da resposta imune adaptativa, tanto humoral quanto mediado por células. Esses epítopos têm grande utilidade em diversas aplicações, como no diagnóstico de doenças, desenvolvimento de fármacos e elaboração de vacinas, entre outras áreas (DE GROOT, 2006).

## 1.9 BIOINFORMÁTICA NA PREDIÇÃO DE EPÍTOPOS LINFOCITÁRIOS

O avanço das ferramentas de biologia computacional tem desempenhado papel central na identificação e caracterização de epítopos linfocitários, especialmente no contexto de patógenos virais emergentes. A predição *in silico* de epítopos de células B e T permite a triagem inicial de regiões antigênicas com maior probabilidade de reconhecimento imunológico, reduzindo significativamente custos e tempo experimental (DE GROOT; RAPPUOLI, 2004).

No caso específico dos epítopos de células B, algoritmos contemporâneos incorporam múltiplos descritores físico-químicos, como hidrofilicidade dos resíduos, acessibilidade ao solvente e flexibilidade estrutural, para prever a propensão antigênica, além de técnicas de aprendizado de máquina e redes neurais profundas. Estudos demonstram que abordagens integrativas apresentam desempenho superior às metodologias baseadas em parâmetros isolados (EL-MANZALAWY; DOBBS; HONAVAR, 2017; ZENG *et al.*, 2023). A incorporação de métodos de inteligência artificial tem permitido melhorar a sensibilidade e especificidade na identificação de epítopos lineares (YU; JOSEPH-MCCARTHY; VAJDA, 2025).

No que se refere à predição de epítopos de células T, os algoritmos contemporâneos se concentram, principalmente, na modelagem da interação entre peptídeos derivados do antígeno e moléculas de MHC das classes I e II. Diferentemente dos epítopos de células B, que dependem majoritariamente da exposição superficial do antígeno intacto, os epítopos de células T correspondem a peptídeos processados intracelularmente e apresentados por moléculas de MHC, sendo o reconhecimento condicionado à afinidade de ligação ao sulco do MHC e à estabilidade do complexo peptídeo–MHC (NIELSEN *et al.*, 2010). Nesse contexto, ferramentas modernas como NetMHC e NetMHCII utilizam redes neurais artificiais treinadas com grandes conjuntos de dados experimentais de afinidade de ligação, permitindo predições quantitativas robustas (JURTZ *et al.*, 2017). Mais recentemente, modelos baseados em *deep learning* e integração de dados de espectrometria de massa – como NetMHCpan – ampliaram a capacidade de predição para múltiplos alelos humanos e não humanos, superando limitações associadas à escassez de dados experimentais para determinados haplótipos (REYNISSON *et al.*, 2020). Esses avanços consolidaram a imunoinformática como ferramenta essencial no mapeamento de epítopos T, especialmente no desenvolvimento de vacinas multiepítopo e na identificação de regiões conservadas capazes de induzir respostas celulares amplas e potencialmente protetoras.

Além disso, bases de dados especializadas têm contribuído para o refinamento dos modelos preditivos ao disponibilizar conjuntos extensivos de epítopos experimentalmente validados (VITA *et al.*, 2019). A integração entre bancos de dados imunológicos e ferramentas de modelagem estrutural amplia a capacidade de análise comparativa entre variantes virais e permite identificar regiões conservadas com potencial imunogênico.

No contexto do desenvolvimento vacinal, a imunoinformática se tornou ferramenta estratégica para o *design* racional de imunógenos, especialmente em abordagens de vacinas baseadas em epítopos. Essa estratégia é particularmente relevante para vírus com alta variabilidade genética ou múltiplos filogrupos, nos quais a proteção cruzada pode ser limitada (DE GROOT; RAPPUOLI, 2004). Assim, a biologia computacional não apenas otimiza a triagem inicial de candidatos antigênicos, mas também orienta decisões experimentais subsequentes, contribuindo para maior eficiência translacional.

# 2 OBJETIVOS

## 2.1 GERAL

- Analisar comparativamente a ocorrência de epítopos lineares de células B e T $CD4^+$ na proteína G de lissavírus dos filogrupos I e II.

## 2.2 ESPECÍFICOS

- Determinar o percentual de identidade entre diferentes versões da proteína G de lissavírus dos filogrupos I e II;
- Estabelecer sequências consenso da proteína G de lissavírus dos filogrupos I e II;
- Predizer a ocorrência de epítopos lineares de células T $CD4^+$ na sequência consenso da proteína G de lissavírus dos filogrupos I e II;
- Predizer a ocorrência de epítopos lineares de células B na sequência consenso da proteína G de lissavírus dos filogrupos I e II.

# 3 MATERIAL E MÉTODOS

## 3.1 ASPECTOS ÉTICOS E LEGAIS

A presente pesquisa, por ser de caráter exclusivamente computacional, não envolve manipulação direta de organismos vivos, tecidos, amostras biológicas humanas ou animais, tampouco procedimentos experimentais que envolvam risco à saúde ou bem-estar de seres vivos. Dessa forma, não houve experimentos em animais ou seres humanos, dispensando, portanto, a apreciação e aprovação por Comitê de Ética no Uso de Animais (CEUA) e Comitê de Ética em Pesquisa com Seres Humanos (CEP).

Por outro lado, a pesquisa foi cadastrada no Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SisGen) sob o número A83F17B (Anexo), uma vez que envolveu atividade de acesso ao patrimônio genético de vírus, disponíveis em bancos de dados públicos. Embora não tenha havido manipulação direta desses agentes infecciosos, esse procedimento garantiu conformidade com as regulamentações nacionais e internacionais sobre a utilização de dados baseados em patrimônio genético, assegurando o alinhamento da pesquisa com as boas práticas éticas e legais no campo da bioinformática.

## 3.2 COLETA DAS SEQUÊNCIAS AMINOACÍDICAS

A seleção de representantes de lissavírus dos filogrupos I e II neste estudo foi feita com base em sua relevância para a saúde pública, considerando sua importância médica e potencial zoonótico associado a eventos de transmissão interespecífica.

A sequência aminoacídica completa da proteína G dos lissavírus membros dos filogrupos I e II foi obtida por tradução conceitual em formato FASTA a partir da sequência de referência do genoma de cada espécie, disponível no banco de dados GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) com os seguintes códigos de acesso: RABV (NP_056796.1), EBLV-1 (YP_001285391.1), EBLV-2 (YP_001285396.1), DUVV (YP_007641405.1), ABLV (NP_478342.1), LBV (YP_007641390.1) e MOKV (YP_142353.1).

Em seguida, utilizaram-se os servidores SignalP 6.0 (TEUFEL *et al.*, 2022) e TMHMM 2.0 (KROGH *et al.,* 2001) para realizar a identificação dos resíduos relativos a peptídeos sinais e hélices transmembrana, respectivamente, da proteína G de cada um dos

membros dos filogrupos I e II, de modo a isolar para análise apenas a sequência aminoacídica relativa ao seu ectodomínio.

## 3.3 DETERMINAÇÃO DA MATRIZ DE IDENTIDADE PERCENTUAL

Para montagem da matriz de identidade percentual, as sequências aminoacídicas passaram por alinhamento múltiplo na ferramenta Clustal Omega 1.2.4, definida para a ordem de entrada dos dados e formato de saída com contagem de caracteres (SIEVERS *et al.*, 2011). Em seguida, as sequências alinhadas foram submetidas à ferramenta EMBOSS Cons 6.6.0 (RICE; LONGDEN; BLEASBY, 2000) para a geração de duas sequências consenso, uma para cada filogrupo de lissavírus.

## 3.4 PREDIÇÃO DE EPÍTOPOS LINEARES DE CÉLULAS T $CD4^+$ E B

Posteriormente, os arquivos FASTA foram submetidos à ferramenta NetMHCIIpan 4.1 (REYNISSON *et al.*, 2020) para a predição de afinidade de peptídeos derivados da proteína G a moléculas HLA-II. Para a submissão, utilizou-se o comprimento de 15 resíduos, selecionando os seguintes alelos: DRB1*03:01, DRB1*07:01, DRB1*15:01, DRB3*01:01, DRB3*02:02, DRB4*01:01 e DRB5*01:01. Na sequência, os dados obtidos foram processados no servidor Epitope-Evaluator (SOTO *et al.*, 2022), considerando o limiar de afinidade de percentil (*rank*) de 5% bem como a localização de cada peptídeo ao longo da estrutura primária do ectodomínio da proteína G.

Para a predição de epítopos de células B, avaliaram-se os perfis de hidrofilicidade, acessibilidade e flexibilidade dos resíduos ao longo do ectodomínio da proteína G dos lissavírus dos filogrupos I e II, utilizando a ferramenta PCProf (COMBET *et al.*, 2000), configurada quanto ao tamanho da janela de análise em 7 aminoácidos. Para permitir a comparação direta entre parâmetros com escalas diferentes, todos os valores foram normalizados para o intervalo de 0 a 1. A pontuação composta foi então calculada como a média aritmética simples dos valores normalizados para cada resíduo, atribuindo contribuição igual à hidrofilicidade, acessibilidade e flexibilidade. Os três resíduos localizados em cada região N-terminal e C-terminal foram excluídos da análise para evitar efeitos de borda inerentes aos cálculos baseados em janelas deslizantes, que resultam em valores artificialmente baixos ou indefinidos nos limites da sequência. Regiões com maior potencial antigênico foram identificadas aplicando-se um limiar correspondente aos 50% superiores das pontuações compostas e considerando-se um mínimo de 5 resíduos

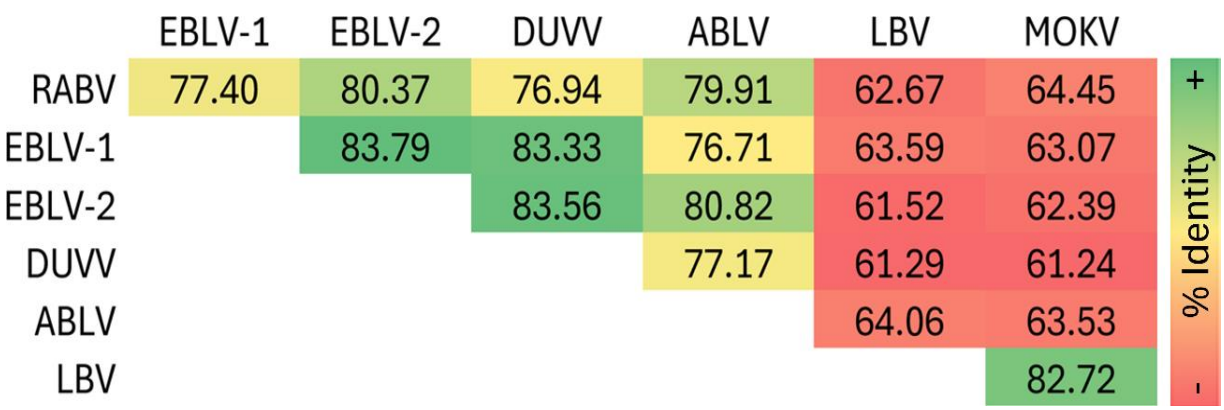
contíguos. Por fim, os epítopos identificados foram checados quanto à chance de mascaramento por N-glicosilação através do servidor NetNGlyc 1.0 (GUPTA & BRUNAK, 2002), que utiliza redes neurais artificiais para examinar a ocorrências de séquons N-X-S/T, considerando-se um limiar de 0,5 para o potencial de modificação pós-traducional.

# 4 RESULTADOS

## 4.1 PERCENTUAL DE IDENTIDADE DAS SEQUÊNCIAS AMINOACÍDICAS

A análise da matriz de identidade percentual do ectodomínio da proteína G revelou níveis moderados a elevados de conservação entre as sequências avaliadas de RABV, EBLV-1, EBLV-2, DUVV, ABLV, LBV e MOKV. Os valores de identidade variaram de 61,24% (entre DUVV e MOKV) a 83,79% (entre EBLV-1 e EBLV-2), indicando que todas as sequências compartilham homologia significativa, compatível com sua classificação dentro do gênero *Lyssavirus* (Figura 6).

**Figura 6:** Matriz de identidade percentual do ectodomínio da proteína G dos lissavírus dos filogrupos I e II.




**Fonte:** Autor.

No contexto do filogrupo I, RABV apresentou menor valor de identidade quando comparado a DUVV (76,94%) e maior valor de identidade quando comparado a EBLV-2 (80,37%). Já no contexto do filogrupo II, a análise da matriz revelou um nível moderado de conservação entre LBV e MOKV, com 82,72% de identidade. As sequências consenso obtidas a partir do alinhamento múltiplo do ectodomínio da proteína G dos lissavírus de cada filogrupo estão representadas na Figura 7, evidenciando 439 resíduos para o filogrupo I e 436 resíduos para o filogrupo II.

**Figura 7:** Sequências consenso do ectodomínio da proteína G dos lissavírus do filogrupo I (PG-I) e do filogrupo II (PG-II).

```
>PG-I
KFPIYTIPDKLGPWSPIDIHHLSCPNNLVVEDEGCTTLTxFSYMELKVGYITTIKVNGFT
CTGVVTEAETYTNFVGYVTTTFKRKHFRPTPSACRDAYNWKIAGDPRYEESLHNPYPDSH
WLRTVKTTKESLLIISPSVADMDAYDKSLYSKMFPNGKCSGVSxxSxFCSTNHDYTIWMP
ENPNPGMSCDIFTxSKGKKATKDGKLCGFVDERGLYKSLKGACKLKLCGISGLRLMDGSW
VSLQxxNExKWCSPDQLVNIHDFHSDEIEHLVVEELVKKREECLDALESIITTKSISFRR
LSHLRKLVPGFGKAYTIINKTLMEADAHYKSVREWxEIIPSKGCLKAGGRCHPHYNGVFF
NGIILGPDGEVLIPEMQSALLQQHIELLESSVIPLMHPLADPSTVFKxDDEAEDFVEVHL
PDVQKQISGIDLGLPEWKx
```

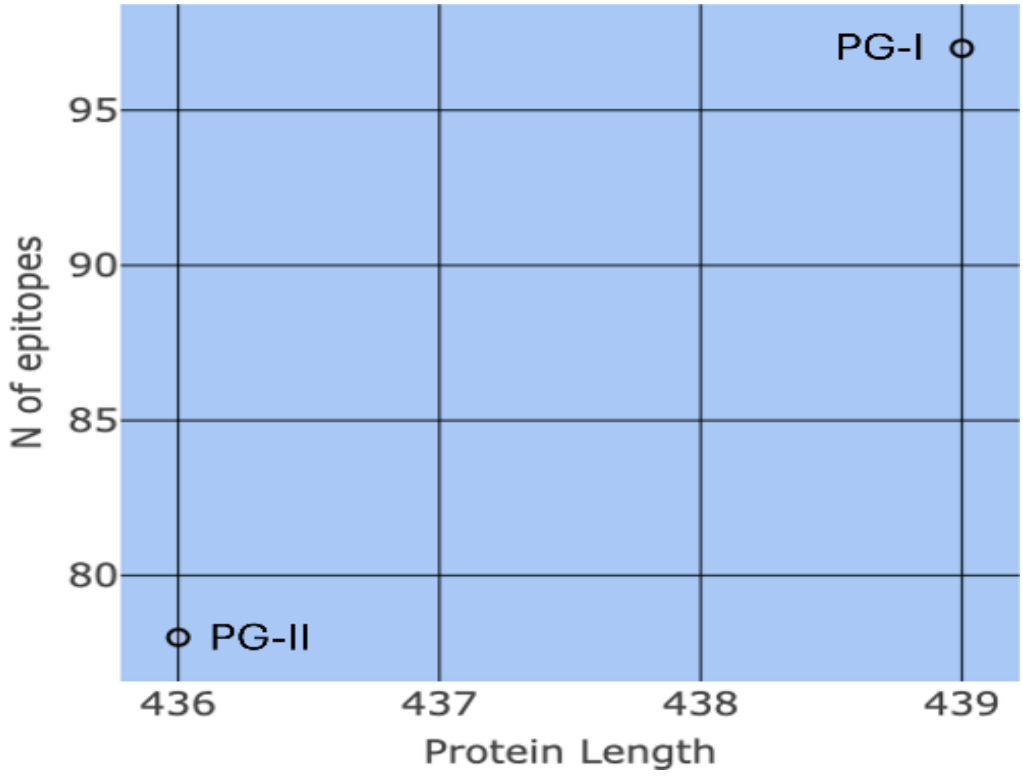
```
>PG-II
DFPLYTIPERLnpWTPIDLIHLSCPNNLLSDaEGCSetSSFTYvELKTGFLAHQKVPGFT
CTGVINEAvTYTNFVGYVTTTFKRKHFKPTVSACRDAFNWKVSGDPRYEESLHTPYPDNS
WLRTVTTTKESLLIISPSIVEMDVYGRTLHSPMFPTGICSkfYPSIPSCaTNHDYTLWLP
EDPNLSLICDIFvTStGRKAMNGSRMCGFtDERGFYRTLKGACKLTLCGKPGLRLYDGTW
VSFTRPEIHVWCSPDQLVNVHNNRIDEIEHLIVDDLIRKREECLDTLETVLMSKSLSFRR
LSHFRKLVPGYGKAYTILNGSLMETNVHYKRVDnWvDILPSKGCLKVnnKCMEsdtGVfF
NGIIKGPDGRILIPEMQSglLKQHMDLLKAAVFPLRHPLIdKaSVFKKDGDADDFVDVHM
PDIqKlVSDVDLGLph
```

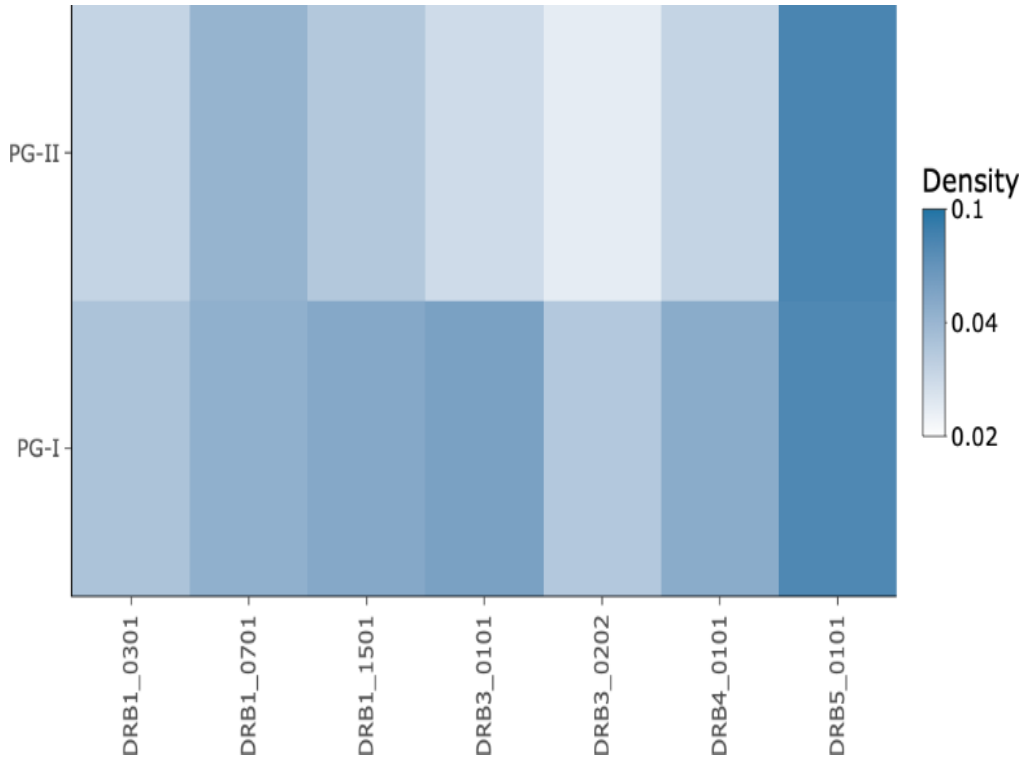
**Fonte:** Autor.

## 4.2 PREDIÇÃO DE EPÍTOPOS LINEARES DE CÉLULAS T CD4$^{+}$

A predição de epítopos para HLA-II na estrutura primária do ectodomínio da proteína G revelou 97 peptídeos para os lissavírus do filogrupo I e 78 peptídeos para os lissavírus do filogrupo II, com uma densidade de 0,22 e 0,18, respectivamente (painel da esquerda na Figura 8). Para ambos os filogrupos, o alelo DRB3*02:02 concentrou a menor densidade de epítopos – 0,03 (14 peptídeos) no filogrupo I e 0,02 (9 peptídeos) no filogrupo II – e o alelo DRB5*01:01 concentrou a maior densidade de epítopos – 0,07 (32 peptídeos) no filogrupo I e 0,08 (33 peptídeos) no filogrupo II (painel da direita na Figura 8).

**Figura 8:** Densidade de epítopos para HLA-II no ectodomínio da proteína G de lissavírus do filogrupo I (PG-I) e do filogrupo II (PG-II).

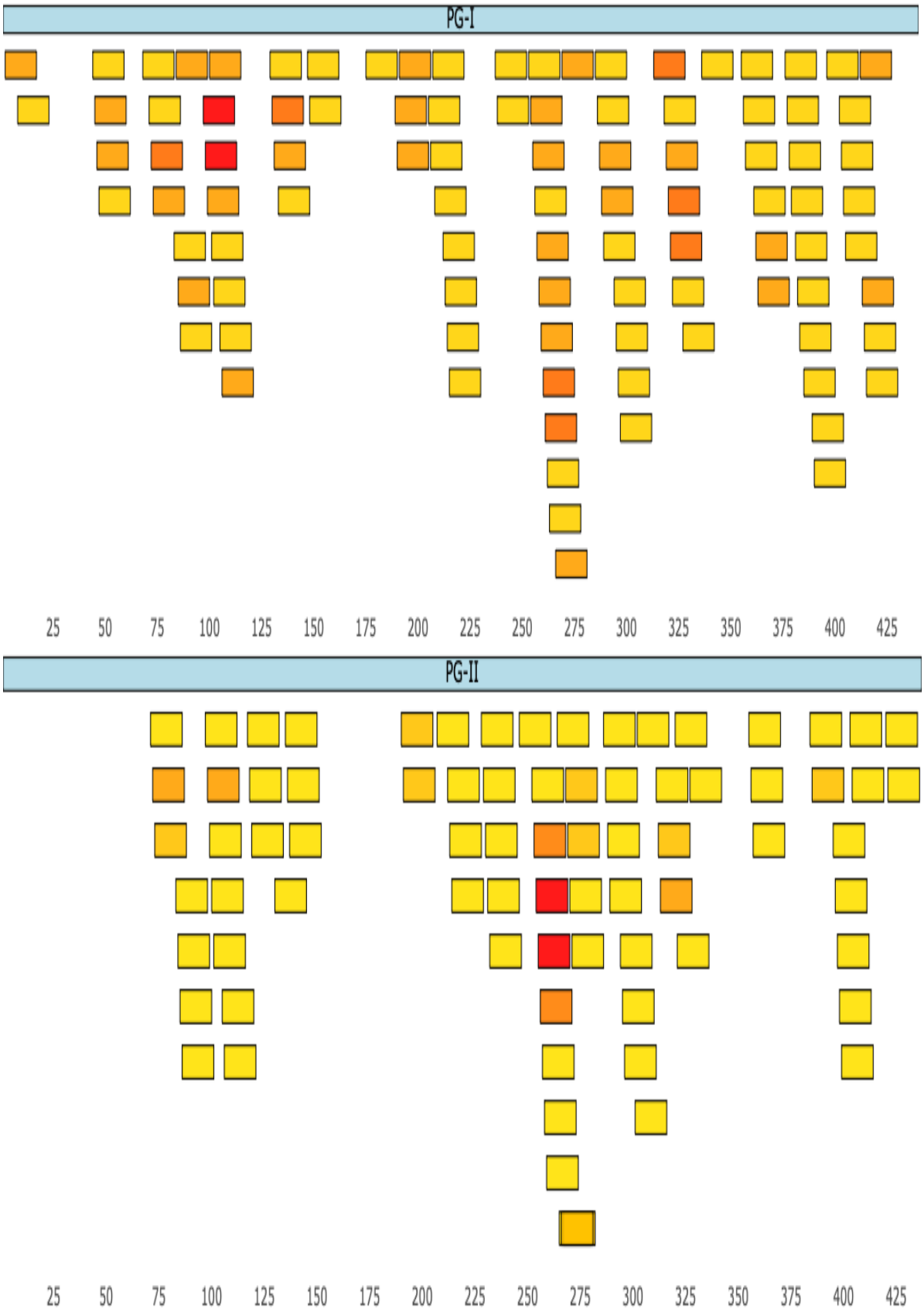





**Fonte:** Autor.

Para os lissavírus do filogrupo I, 2 peptídeos sobrepostos com afinidade por 4 alelos de HLA-II foram reconhecidos. O primeiro, alocado na posição 96–110, foi identificado pela sequência aminoacídica DAYNWKIAGDPRYEE, enquanto o segundo, alocado na posição 97–111, foi identificado pela sequência aminoacídica AYNWKIAGDPRYEES. Ambos os peptídeos demonstraram afinidade pelos mesmos alelos: DRB1*03:01, DRB3*01:01, DRB3*02:02 e DRB5*01:01 (caixas em vermelho no painel superior da Figura 9). Já para os lissavírus do filogrupo II, 2 peptídeos sobrepostos com afinidade por 6 alelos de HLA-II foram reconhecidos. O primeiro, alocado na posição 253–267, foi identificado pela sequência aminoacídica SPDQLVNVHNNRIDE, ao passo que o segundo, alocado na posição 254–268, foi identificado pela sequência aminoacídica PDQLVNVHNNRIDEI. Ambos os peptídeos também demonstraram afinidade pelos mesmos alelos: DRB1*07:01, DRB1*15:01, DRB3*01:01, DRB3*02:02, DRB4*01:01 e DRB5*01:01 (caixas em vermelho no painel inferior da Figura 9).

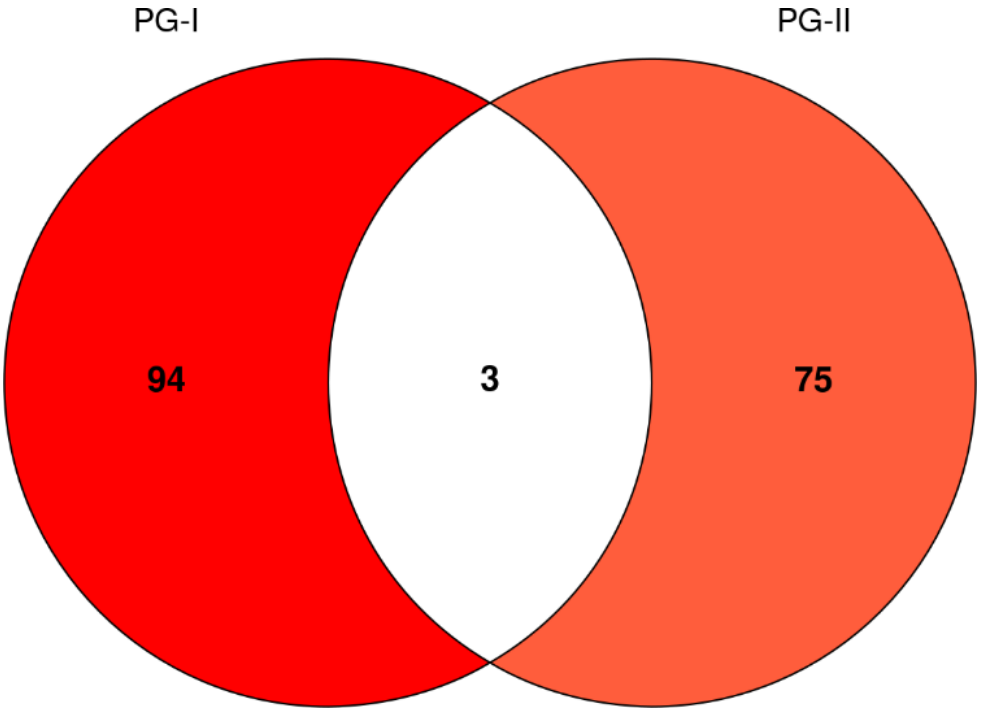
**Figura 9:** Predição de epítopos lineares de células T para HLA-II na estrutura primária (em azul) do ectodomínio da proteína G de lissavírus do filogrupo I (PG-I) e do filogrupo II (PG-II), coloridos de acordo com a promiscuidade alélica em um gradiente de amarelo (mínima) a vermelho (máxima).




**Fonte:** Autor.

A análise da conservação de epítopos entre os filogrupos de lissavírus revelou que, embora a maior parte dos peptídeos identificados fossem exclusivos dos filogrupos I (94) e II (75), 3 peptídeos sobrepostos foram compartilhados entre ambos os filogrupos: TYTNFVGYVTTTFKR (posição 70–84; alelo DRB1*15:01), YTNFVGYVTTTFKRK (posição 71–85; alelos DRB1*07:01, DRB1*15:01 e DRB5*01:01) e TNFVGYVTTTFKRKH (posição 72–86; alelos DRB1*15:01 e DRB5*01:01) (Figura 10).

**Figura 10:** Conservação de epítopos para HLA-II no ectodomínio da proteína G de lissavírus do filogrupo I (PG-I) e do filogrupo II (PG-II).



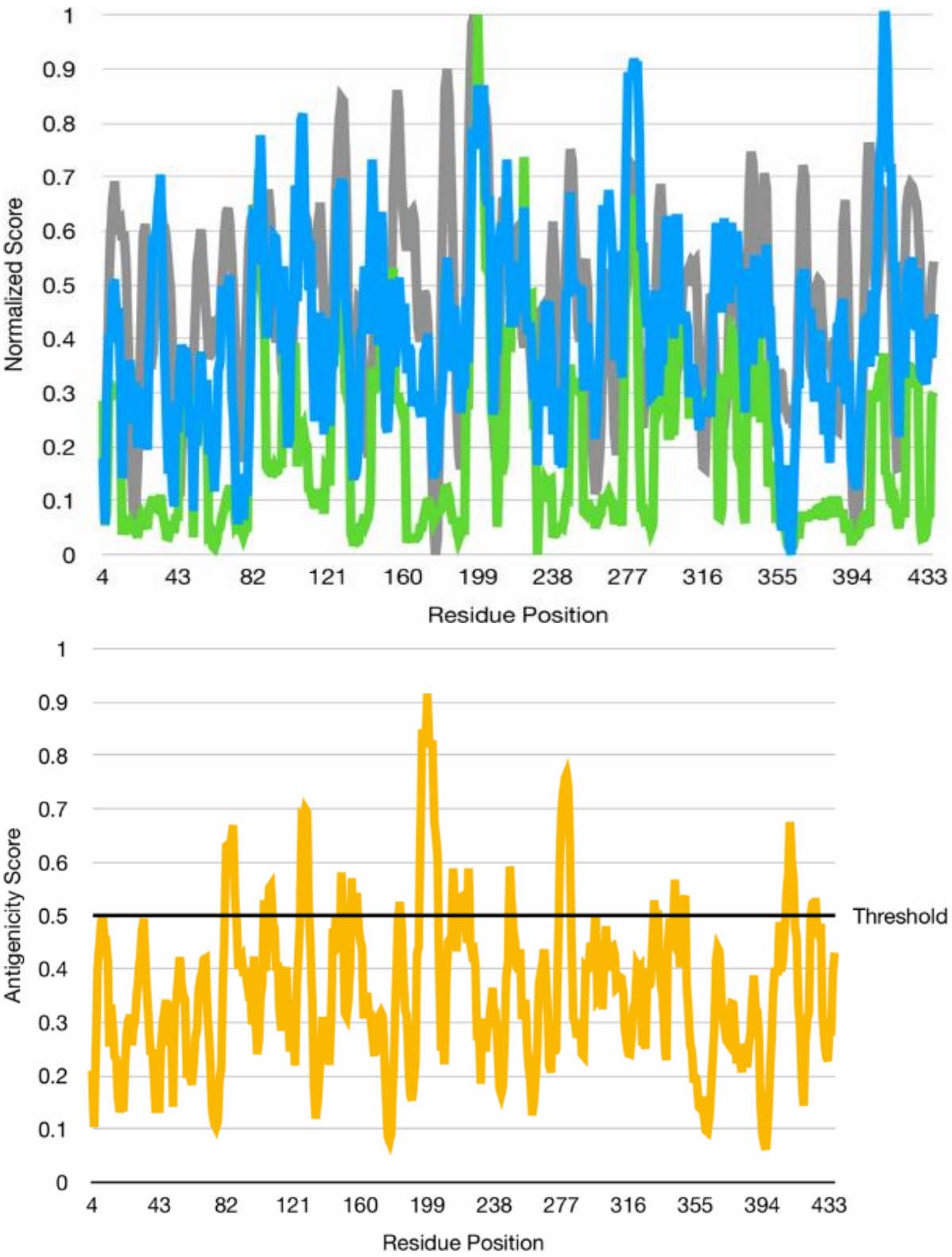

**Fonte:** Autor.

## 4.3 PREDIÇÃO DE EPÍTOPOS LINEARES DE CÉLULAS B

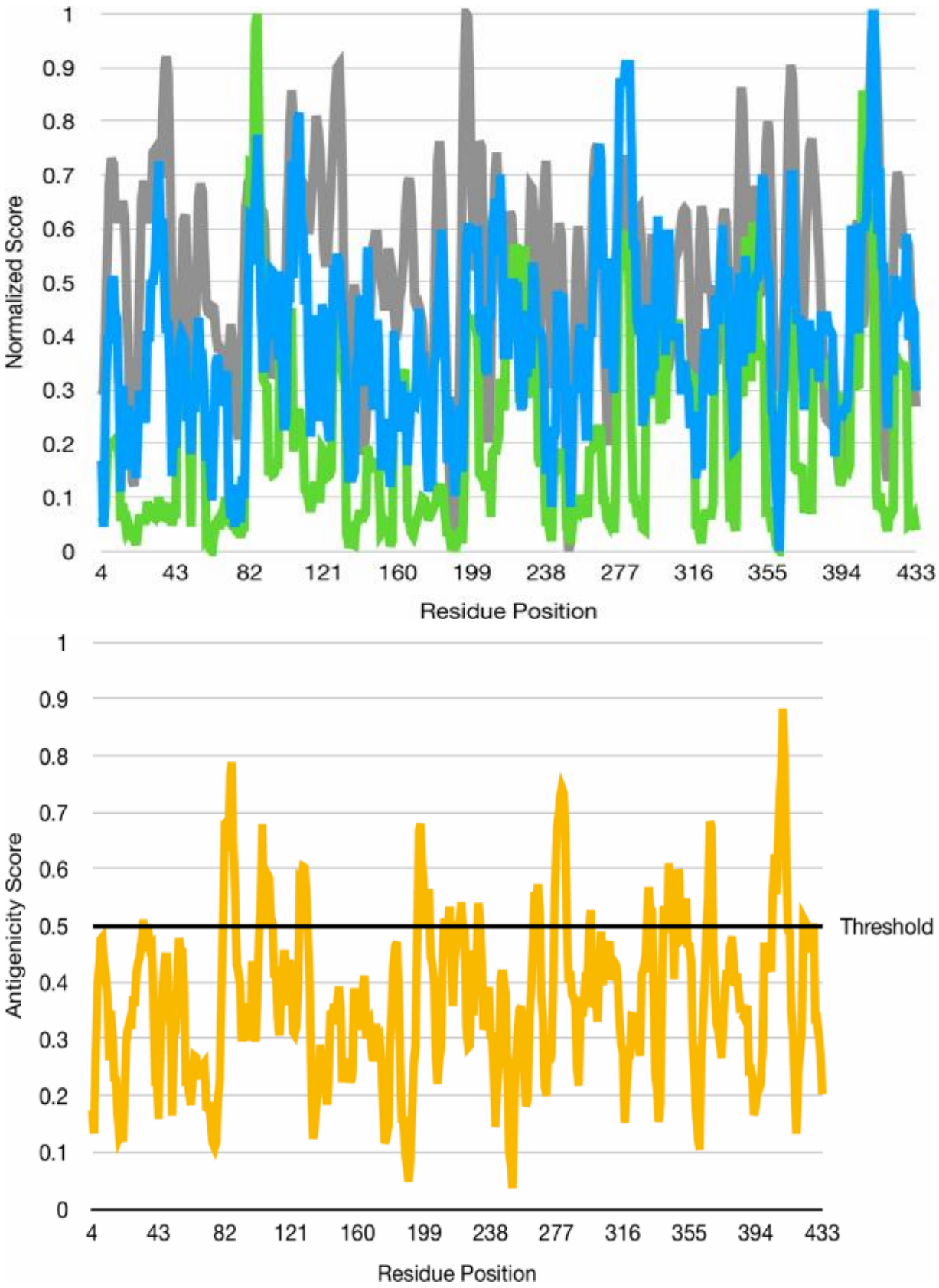
Nos lissavírus do filogrupo I, a antigenicidade média foi maior na metade N-terminal do ectodomínio da proteína G, alcançando o limiar nas posições 82–87 (FKRKHF), 125–130 (VKTTKE), 195–205 (SKGKKATKDGK), 217–221 (KSLKG), 276–282 (LVKKREE) e 408–412 (XDDEA), enquanto, nos lissavírus do filogrupo II, a antigenicidade média foi maior na metade C-terminal do ectodomínio da proteína G, atingindo o limiar nas posições 81–88 (TFKRKHFK), 106–110 (PRYEE), 126–130 (TTTKE), 196–202 (TGRKAMN), 276–283 (LIRKREEC) e 404–412 (SVFKKDGDA) (Figura 11).

**Figura 11:** Antigenicidade (em amarelo) no ectodomínio da proteína G de lissavírus do filogrupo I (PG-I) e do filogrupo II (PG-II) calculada a partir dos parâmetros de hidrofilicidade (em azul), acessibilidade (em verde) e flexibilidade (em cinza).

PG-I



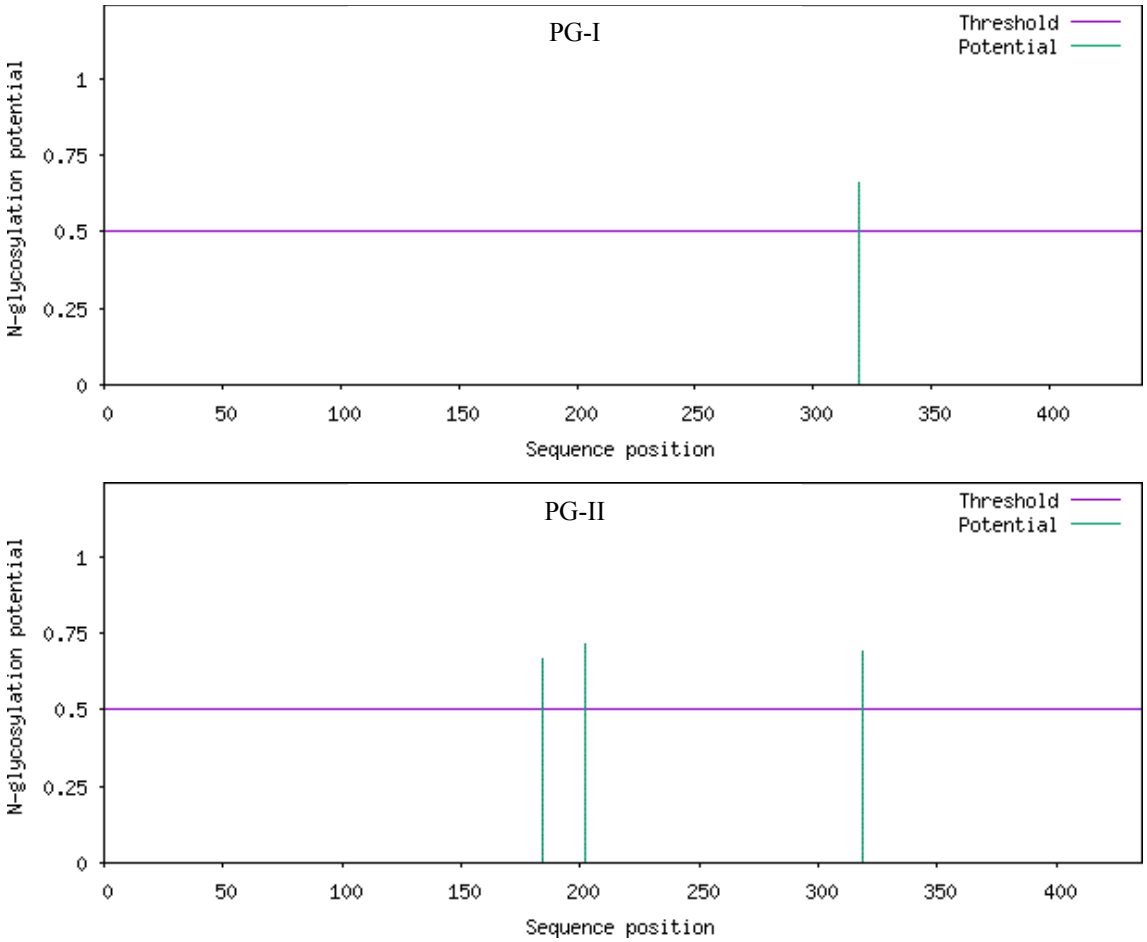
PG-II




**Fonte:** Autor

Quanto ao potencial de N-glicosilação no ectodomínio da proteína G, identificou-se a ocorrência de 1 séquon nos lissavírus do filogrupo I – na posição 319 (NKTL) – e 3 séquons nos lissavírus do filogrupo II – nas posições 184 (NLSL), 202 (NGSR) e 319 (NGSL) (Figura 12).

**Figura 12:** Potencial de N-glicosilação no ectodomínio da proteína G de lissavírus do filogrupo I (PG-I) e do filogrupo II (PG-II).




**Fonte:** Autor

## 5 DISCUSSÃO

A predição *in silico* de epítopos restritos ao HLA-II no ectodomínio da proteína G dos lissavírus dos filogrupos I e II revelou que os peptídeos de maior promiscuidade alélica estão localizados em posições distintas na molécula (96–110/97–111 e 253–267/254–268, respectivamente). Ainda assim, tais peptídeos compartilharam afinidade de ligação com múltiplos alelos de HLA-II, incluindo DRB3*01:01, DRB3*02:02 e DRB5*01:01, conforme evidenciado pelos algoritmos de predição utilizados. Adicionalmente, a densidade de epítopos para HLA-II foi ligeiramente maior no filogrupo I (0,22) do que no filogrupo II (0,18), indicando que a proteína G do primeiro filogrupo

constitui um alvo relativamente mais imunogênico para respostas mediadas por células T $CD4^+$.

A identificação de epítopos sobrepostos com afinidade por um conjunto diversificado de alelos de HLA-II é um achado particularmente relevante, pois regiões proteicas contendo tais epítopos tendem a favorecer uma apresentação antigênica mais eficiente, aumentando a probabilidade de ativação de células T $CD4^+$ e, consequentemente, potencializando a resposta imune adaptativa humoral e celular (ABBAS; LICHTMAN; PILLAI, 2021). Ademais, os alelos observados são frequentemente relatados na literatura como participantes centrais na apresentação de epítopos derivados de vírus de importância médica, sugerindo que os lissavírus dos filogrupos I e II compartilham determinantes imunológicos capazes de serem reconhecidos por vias clássicas da imunidade adaptativa humana (CASTELLI *et al.*, 2021; WEISKOPF *et al.*, 2013).

O reconhecimento de regiões candidatas a epítopos lineares de células B no ectodomínio da proteína G dos lissavírus dos filogrupos I e II foi fundamentado na análise combinada de hidrofilicidade, acessibilidade e flexibilidade como *proxy* para antigenicidade. Enquanto regiões hidrofílicas apresentam maior probabilidade de estarem localizadas na superfície proteica e, portanto, de serem reconhecidas por anticorpos (ZENG *et al.*, 2023), a acessibilidade é considerada um requisito estrutural essencial para o reconhecimento por células B, uma vez que apenas resíduos expostos podem participar da formação do complexo antígeno-anticorpo (EL-MANZALAWY; DOBBS; HONAVAR, 2017). Além disso, epítopos lineares frequentemente se localizam em regiões estruturalmente permissivas e com maior mobilidade conformacional (ZENG *et al.*, 2023); tal flexibilidade estrutural, por sua vez, contribui para a capacidade adaptativa do epítopo à ligação com o anticorpo (YU; JOSEPH-MCCARTHY; VAJDA, 2025).

Embora a antigenicidade média tenha sido maior em porções distintas do ectodomínio da proteína G em cada filogrupo de lissavírus (i.e., metade N-terminal para o filogrupo I e metade C-terminal para o filogrupo II), apenas 1 dos 6 epítopos lineares de células B que superaram o limiar em cada filogrupo não apresentou sobreposição de sequência (posições 217–221 no filogrupo I e 106–110 no filogrupo II), sugerindo uma propensão diferencial de ambos os filogrupos à iniciação de respostas adaptativas humorais.

Dos distintos sítios antigênicos já reconhecidos em estudos prévios sobre a imunogenicidade da proteína G do RABV, apenas o sítio antigênico 1 (AS1) é um epítopo linear – os outros sítios antigênicos (AS2, AS3, AS4 e ASa) são epítopos conformacionais e não contínuos (SHI *et al.*, 2022). O AS1 foi caracterizado utilizando um anticorpo monoclonal (CR57) capaz de se ligar aos resíduos KLCGVL (posição 226–231) (MARISSEN *et al.*, 2005). No presente estudo, tal posição é representada pelos resíduos KLCGIS no filogrupo I e TLCGKP no filogrupo II, mas não superou o limiar de antigenicidade estabelecido na análise. No entanto, as posições 195–205 (SKGKKATKDGK) do filogrupo I e 196–202 (TGRKAMN) do filogrupo II, que superaram tal limiar, sobrepõem-se à posição 198–200 (KRA) do AS2, composto também pela posição 34–42 (GCTNLSGFS) (NAGARAJAN; MARISSEN; RUPPRECHT, 2014). Apesar da sobreposição, não há conservação total da sequência reconhecida em ambos os casos.

Vale destacar que os epítopos lineares sobrepostos de células B contidos nas posições 82–87 (FKRKHF) do filogrupo I e 81–88 (TFKRKHFK) do filogrupo II também se sobrepuseram em, pelo menos, 3 resíduos (FKR) aos peptídeos TYTNFVGYVTTTFKR (posição 70–84), YTNFVGYVTTTFKRK (posição 71–85) e TNFVGYVTTTFKRKH (posição 72–86), comumente ligados pelo alelo DRB1*15:01 de HLA-II, sugerindo relevância funcional em respostas de anticorpos dependentes de células T $CD4^+$. Ademais, tais epítopos não foram mascarados pela presença de N-glicanos, sugerindo que esta região do ectodomínio da proteína G deve ser priorizada no desenvolvimento de imunobiológicos efetivos simultaneamente contra ambos os filogrupos de lissavírus.

## 6 CONCLUSÃO

O presente estudo alcançou o objetivo de analisar comparativamente a ocorrência de epítopos lineares de células B e T CD4+ na proteína G de lissavírus pertencentes aos filogrupos I e II, contribuindo para o entendimento do perfil imunogênico desses vírus. Ademais, foi possível determinar o percentual de identidade entre diferentes versões da proteína G, evidenciando níveis moderados a elevados de conservação entre as sequências analisadas. No filogrupo I, os valores de identidade variaram de 76,71% a 83,79%, enquanto no filogrupo II foi observada identidade de 82,72%, demonstrando que, apesar da variabilidade interespecífica, há manutenção de regiões conservadas potencialmente relevantes do ponto de vista antigênico.

Em consonância com o segundo objetivo específico, foram estabelecidas sequências consenso do ectodomínio da proteína G para os filogrupos I e II, com 439 e 436 resíduos aminoacídicos, respectivamente. A obtenção dessas sequências permitiu padronizar a análise comparativa entre os filogrupos e constituiu a base para a investigação subsequente da distribuição de epítopos lineares, possibilitando uma abordagem integrada das propriedades imunogênicas da proteína G.

Quanto à predição de epítopos lineares de células T CD4+, observou-se maior densidade de peptídeos com afinidade por moléculas de HLA-II no filogrupo I em comparação ao filogrupo II. Também foram identificados peptídeos de elevada promiscuidade alélica e três epítopos conservados compartilhados entre ambos os filogrupos, sugerindo a existência de determinantes antigênicos potencialmente relevantes para o reconhecimento imune mediado por células T auxiliares. Esses achados indicam que, embora existam diferenças no perfil de imunogenicidade entre os filogrupos, regiões conservadas podem representar alvos importantes para respostas celulares amplas.

Em relação aos epítopos lineares de células B, a análise revelou distribuição diferencial da antigenicidade ao longo do ectodomínio da proteína G, com predominância na metade N-terminal no filogrupo I e na metade C-terminal no filogrupo II. Destaca-se que epítopos localizados em posições equivalentes entre os dois filogrupos apresentaram sobreposição com peptídeos reconhecidos pelo alelo HLA-II DRB1*15:01 e não demonstraram potencial de mascaramento por N-glicosilação. Em conjunto, esses resultados indicam regiões com potencial relevância funcional para respostas humorais dependentes de células T CD4+, fornecendo subsídios para estudos experimentais futuros e para o

desenvolvimento de estratégias vacinais e imunoprofiláticas mais abrangentes contra lissavírus.

## 7 BIBLIOGRAFIA

ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. *Cellular and Molecular Immunology*. 10ª ed. Philadelphia: Elsevier, 2021.

ÁVILA, R. A. M. Predição de epítopos descontínuos ou conformacionais em proteínas através da bioinformática estrutural. Dissertação (Pós-Graduação em Bioinformática) – Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

BADRANE, H.; TORDO, N. Mudança de hospedeiro na história do lyssavirus das ordens Chiroptera para Carnivora. *Journal of Virology*, v. 75, p. 8096–8104, 2001.

BANYARD, A. C. et al. Lyssaviruses and bats: emergence and zoonotic threat. *Viruses*, v. 5, n. 12, p. 2974–2990, 2013.

BATISTA, H. B. C. R.; FRANCO, C. A.; ROEHE, P. M. Raiva: uma breve revisão. *Acta Scientiae Veterinariae*, v. 35, n. 2, p. 125–144, 2007.

BEN HAMED, S. et al. Toward the development of a pan-lyssavirus vaccine. *Viruses*, v. 16, n. 7, p. 1107, 2024.

BOULGER, L. R.; PORTERFIELD, J. S. Isolation of a virus from Nigerian fruit bats. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, v. 52, p. 421–424, 1958.

BROOKES S.M, et al. Rabies human diploid cell vaccine elicits cross-neutralising and cross-protecting immune responses against European and Australian bat lyssaviruses. *Vaccine*, 23(32):4101–4109, 2005.

CALLAWAY, H. M. et al. Structure of the rabies virus glycoprotein trimer bound to a prefusion-specific neutralizing antibody. *Science Advances*, v. 8, p. eabp9151, 2022.

CASTELLI, E. C. et al. MHC variants associated with symptomatic versus asymptomatic SARS-CoV-2 infection in highly exposed individuals. *Frontiers in Immunology*, v. 12, p. 742881, 2021.

COERTSE, J. et al. Lagos bat virus: an under-reported rabies-related lyssavirus. *Viruses*, v. 13, n. 4, p. 576, 2021.

COMBET, C. et al. NPS@: network protein sequence analysis. *Trends in Biochemical Sciences*, v. 25, n. 3, p. 147–150, 2000.

DAVIS, B. M.; RALL, G. F.; SCHNELL, M. J. Everything you always wanted to know about rabies virus (but were afraid to ask). *Annual Review of Virology*, v. 2, n. 1, p. 451–471, 2015.

DE GROOT, A. S. Immunomics: discovering new targets for vaccines and therapeutics. *Drug Discovery Today*, v. 11, n. 5-6, p. 203–209, 2006.

DE GROOT, A. S.; RAPPUOLI, R. Genome-derived vaccines. *Expert Review of Vaccines*, v. 3, n. 1, p. 59–76, 2004.

DHULIPALA, S.; UVERSKY, V. N. Looking at the pathogenesis of the Rabies lyssavirus strain Pasteur vaccins through a prism of the disorder-based bioinformatics. *Biomolecules*, v. 12, n. 10, p. 1436, 2022.

DURRER, P. et al. Identification of a domain in the rabies virus glycoprotein involved in membrane fusion. *Journal of General Virology*, v. 76, n. 2, p. 315–320, 1995.

ECHEVARRÍA J. E et al. Current Rabies Vaccines Do Not Confer Protective Immunity against Some of the Newly Identified Lyssaviruses. *Viruses*, 11(10):892, 2019.

EL-MANZALAWY, Y.; DOBBS, D.; HONAVAR, V. G. In silico prediction of linear B-cell epitopes on proteins. *Methods in Molecular Biology*, v. 1484, p. 255–264, 2017.

EVANS, J. S. et al. Rabies virus vaccines: Is there a need for a pan-lyssavirus vaccine? *Vaccine*, 30(52):7447–7454, 2012.

EVANS, J. S. et al. Antigenic site changes in the rabies virus glycoprotein. *Journal of General Virology*, v. 99, p. 169–180, 2018.

FABER, M. et al. Overexpression of the rabies virus glycoprotein results in enhancement of apoptosis and antiviral immune response. *Journal of Virology*, v. 83, n. 11, p. 5615–5624, 2009.

FOOKS, A. R. et al. European bat lyssaviruses: an emerging zoonosis. *Epidemiology and Infection*, v. 131, n. 3, p. 1029–1039, 2003.

FRANCIS, J. R. et al. Australian bat lyssavirus: implications for public health. *Medical Journal of Australia*, v. 201, p. 647–649, 2014.

GUPTA, R.; BRUNAK, S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing*, v. 7, p. 310–322, 2002.

HASSAN, B. et al. Evidence of two lyssavirus phylogroups with distinct pathogenicity and immunogenicity. *Journal of Virology*, v. 75, p. 3268–3276, 2001.

JOHNSON, N. et al. Experimental study of European bat lyssavirus type 2 infection. *Journal of General Virology*, v. 89, p. 2662–2672, 2008.

JURTZ, V. et al. NetMHCpan-4.0: improved peptide-MHC interaction predictions. *Journal of Immunology*, v. 199, n. 9, p. 3360–3368, 2017.

KING, A.; DAVIES, P.; LAWRIE, A. The rabies viruses of bats. *Veterinary Microbiology*, v. 23, n. 1–4, p. 165–174, 1990.

KROGH, A. et al. Predicting transmembrane protein topology with a hidden Markov model. *Journal of Molecular Biology*, v. 305, n. 3, p. 567–580, 2001.

LIAN, M.; HUEFFER, K.; WELTZIN, M. Interactions between the rabies virus and nicotinic acetylcholine receptors. *Heliyon*, 2022.

MARISSEN, W. E. et al. Novel rabies virus-neutralizing epitope recognized by human monoclonal antibody: fine mapping and escape mutant analysis. *Journal of Virology*, v. 79, n. 8, p. 4672–4678, 2005.

MCELHINNEY, L. M. et al. Molecular epidemiology and evolution of European bat lyssavirus 2. *International Journal of Molecular Sciences*, v. 19, n. 1, p. 156, 2018.

MCMAHON, W. C. et al. Surveillance of Mokola virus in small mammals. *Onderstepoort Journal of Veterinary Research*, v. 88, n. 1, 2021.

MEDHASI, S.; CHANTRATITA, N. Human leukocyte antigen (HLA) system: genetics and association with bacterial and viral infections. *Journal of Immunology Research*, v. 2022, p. 9710376, 2022.

MENDEL, I. et al. Delineation of the minimal encephalitogenic epitope. *European Journal of Immunology*, v. 26, n. 10, p. 2470–2479, 1996.

NAGARAJAN, T.; MARISSEN, W. E.; RUPPRECHT, C. E. Monoclonal antibodies for the prevention of rabies: theory and clinical practice. *Antibody Technological Journal*, v. 4, p. 1–12, 2014.

NIELSEN, M. et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS One*, v. 5, n. 8, P. e11083, 2010.

PAWESKA, J. T. et al. Fatal human infection with Duvenhage virus. *Emerging Infectious Diseases*, v. 12, p. 1965–1967, 2006.

REYNISSON, B. et al. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation. *Nucleic Acids Research*, v. 48, n. W1, p. W449–W454, 2020.

REZENDE, J. M. et al. História da raiva e sua importância na saúde pública. *Revista da Sociedade Brasileira de Medicina Tropical*, v. 30, n. 1, p. 85–95, 1997.

RICE, P.; LONGDEN, I.; BLEASBY, A. EMBOSS. *Trends in Genetics*, v. 16, n. 6, p. 276–277, 2000.

ROINE, R. O. et al. Rabies encephalitis in Europe: a rare case of human infection with a bat lyssavirus. *Journal of Neurology, Neurosurgery and Psychiatry*, v. 51, n. 9, p. 1110–1112, 1988.

RUPPRECHT, C. E.; HANLON, C. A.; HEMACHUDHA, T. Rabies re-examined. *The Lancet Infectious Diseases*, v. 2, n. 6, p. 327–343, 2002.

SCHNEIDER, M. C.; SANTOS-BURGOA, C. A historical review of the treatment of human rabies. *Revista de Saúde Pública*, v. 28, n. 6, p. 454–463, 1994.

SCOTT, T. P.; NEL, L. H. Lyssaviruses and the fatal encephalitic disease rabies. *Frontiers in Immunology*, v. 12, p. 786953, 2021.

SERVAT, A et al. Cross-Protection of Inactivated Rabies Vaccines for Veterinary Use against Bat Lyssaviruses Occurring in Europe. *Viruses*, 11(10):936, 2019.

SHI, C. et al. Research progress on neutralizing epitopes and antibodies for the rabies virus. *Infectious Medicine*, v. 1, n. 2, p. 262–271, 2022.

SIEVERS, F. et al. Clustal Omega. *Molecular Systems Biology*, v. 7, p. 539, 2011.

SILVA-PINTO, A.; SANTOS, M. L. The role of major histocompatibility complex in the progression of HIV infection. *Acta Médica Portuguesa*, v. 24, s. 4, p. 923–928, 2011.

SOTO, L. F. et al. Epitope-Evaluator. *PLoS One*, v. 17, n. 8, e0273577, 2022.

TEUFEL, F. et al. SignalP 6.0. *Nature Biotechnology*, v. 40, p. 1023–1025, 2022.

VAN THIEL, P. P. A. M. et al. Fatal human rabies due to Duvenhage virus from a bat in Kenya: failure of treatment with coma-induction, ketamine, and antiviral drugs. *PLoS Neglected Tropical Diseases*, v. 3, n. 7, p. e428, 2009.

VITA, R. et al. The Immune Epitope Database (IEDB). *Nucleic Acids Research*, v. 47, p. D339–D343, 2019.

WALKER, P. J. et al. ICTV Virus Taxonomy Profile: Rhabdoviridae. *Journal of General Virology*, v. 99, n. 4, p. 447–448, 2018.

WEISKOPF, D. et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for $CD8^+$ T cells. *Proceedings of the National Academy of Sciences of the United States of America*, v. 110, n. 22, p. E2046–E2053, 2013.

YU, H.; JOSEPH-MCCARTHY, D.; VAJDA, S. Improving B-cell epitope prediction. *Drug Discovery Today*, v. 30, n. 11, 2025.

ZENG, X. et al. Recent progress in antibody epitope prediction. *Antibodies*, v. 12, n. 3, p. 52, 2023.

# 8 ANEXO

**Ministério do Meio Ambiente**
**CONSELHO DE GESTÃO DO PATRIMÔNIO GENÉTICO**
**SISTEMA NACIONAL DE GESTÃO DO PATRIMÔNIO GENÉTICO E DO CONHECIMENTO TRADICIONAL ASSOCIADO**

**Comprovante de Cadastro de Acesso**
**Cadastro nº A83F17B**

A atividade de acesso ao Patrimônio Genético, nos termos abaixo resumida, foi cadastrada no SisGen, em atendimento ao previsto na Lei nº 13.123/2015 e seus regulamentos.

| | |
|---|---|
| Número do cadastro: | **A83F17B** |
| Usuário: | **Carlos Alberto Marques de Carvalho** |
| CPF/CNPJ: | **115.522.647-00** |
| Objeto do Acesso: | **Patrimônio Genético** |
| Finalidade do Acesso: | **Pesquisa** |

**Espécie**

**Lyssavirus rabies**
**Lyssavirus hamburg**
**Lyssavirus helsinki**
**Lyssavirus duvenhage**
**Lyssavirus australis**
**Lyssavirus lagos**
**Lyssavirus mokola**

| | |
|---|---|
| Título da Atividade: | **Análise Comparativa da Ocorrência de Epítopos Lineares na Glicoproteína de Lissavírus dos Filogrupos I e II** |

**Equipe**

| | |
|---|---|
| **Carlos Alberto Marques de Carvalho** | **UEPA** |
| **André Miller Cortêz Lima** | **UEPA** |
| **Taciana Fernandes Souza Barbosa Coelho** | **IEC** |

| | |
|---|---|
| Data do Cadastro: | **06/04/2026 02:00:20** |
| Situação do Cadastro: | **Concluído** |

Conselho de Gestão do Patrimônio Genético
Situação cadastral conforme consulta ao SisGen em **2:00** de **06/04/2026**.

SISTEMA NACIONAL DE GESTÃO DO PATRIMÔNIO GENÉTICO E DO CONHECIMENTO TRADICIONAL ASSOCIADO - SISGEN