



ESCOLA BAHIANA DE MEDICINA E SAÚDE PÚBLICA

Mariana Maciel Portela

**PREDIÇÃO IN SILICO DE EPÍTOPOS DE LINFÓCITOS T CITOTÓXICOS DA
VARIANTE DELTA DO SARS-COV-2**

Salvador - Bahia

2022

MARIANA MACIEL PORTELA

**PREDIÇÃO IN SILICO DE EPÍTOPOS DE LINFÓCITOS T CITOTÓXICOS DA
VARIANTE DELTA DO SARS-COV-2**

Trabalho de Conclusão de Cursos, apresentado ao Curso de Graduação em Medicina da Escola Bahiana de Medicina e Saúde Pública, para aprovação parcial no 4º ano de Medicina.

Orientadora: Dra. Maria Fernanda Rios Grassi

Coorientador: Carlos Gustavo Regis da Silva

Salvador – Bahia

2022

AGRADECIMENTOS

Agradeço à minha mãe, Margareth, por ser minha inspiração na medicina e por tornar tudo isso possível, sempre acreditando em mim e não medindo esforços para me ajudar a atingir meus objetivos.

Aos meus avós, Orlando e Dolores, por todo amor, carinho e atenção em toda minha jornada. Todo meu conhecimento sobre justiça e caráter vem de vocês. Obrigada por estarem presentes e acreditarem em mim até quando eu mesma duvidei.

Ao meu pai, Marcelo, por ser exemplo de força e suporte em todos os momentos.

À minha irmã, Marina, por estar ao meu lado em todos os momentos. Por sempre saber o que dizer e por ser minha eterna companheira.

Aos meus irmãos, Maria e Enzo e aos meus primos, João Pedro e João Pedro por serem motivo de grande alegria em minha vida.

À toda a minha família, por ser tão presente e amorosa em todos os momentos.

Ao meu namorado, José Victor, por ser meu parceiro para todas as horas e por sempre acreditar em mim. Obrigada por sempre tornar meus dias felizes mais felizes ainda e por ser alegria até nos dias difíceis.

Às minhas amigas, Beatriz, Hayala, Mariana e Rebeca, meus presentes que levarei para vida inteira, por fazerem parte dessa caminhada tornando-a leve, alegre e divertida. Tudo com vocês é incrível.

À minha orientadora e ao meu coorientador, professora Maria Fernanda Grassi e professor Carlos Regis, sempre com tanta paciência e zelo para me direcionar na execução deste trabalho.

À minha amiga, Greice, por estar tão presente em todos os momentos, me ajudando, ensinando e colaborando grandemente com este trabalho.

Ao meu professor, Diego, por se mostrar presente e por tanta atenção e preocupação em ensinar os aspectos metodológicos do trabalho.

RESUMO

Portela, MM. Predição *in silico* de epítomos de linfócitos t citotóxicos da variante delta do sars-cov-2 [trabalho de conclusão de curso]. Salvador: Curso de Medicina, Escola Bahiana de Medicina e Saúde Pública, Bahia; 2022.

INTRODUÇÃO: Em janeiro de 2020, cientistas chineses isolaram o SARS-CoV-2, vírus causador da COVID-19, causador de uma crise de saúde mundial que levou a mais de 6 milhões de óbitos. Nesse contexto, surgiram variantes desse mesmo vírus, com destaque para a variante Delta, que apresentou aumento da patogenicidade e transmissibilidade da doença. Entretanto, até os dias atuais, poucos estudos descreveram a resposta imune desencadeado por células T CD8 citotóxicas para o SARS-CoV-2. Dessa forma, é importante identificar os epítomos T potenciais da variante Delta do SARS-CoV-2, uma vez que permitirá inferir sobre a possibilidade de escape da imunidade celular gerada pela primária infecção ou pela vacinação e qual impacto esse escape é capaz de gerar na saúde dos indivíduos. **OBJETIVO:** Predizer *in silico* epítomos da proteína Spike de linfócitos CTL da variante Delta do SARS-CoV-2, comparando-os com epítomos da sequência-mãe isolada de Wuhan. **METODOLOGIA:** Para a predição, foram utilizadas as moléculas de HLA com maior frequência na população brasileira, encontradas no site “The Allele Frequency”. As sequências de nucleotídeos da sequência-referência do SARS-CoV-2 e da variante Delta do SARS-CoV-2 foram obtidas através da plataforma GISAID. A sequência de nucleotídeos da sequência-referência foi traduzida com auxílio do software “Expase Translate” e foram coletadas as sequências de aminoácidos correspondentes à proteína Spike. Após a seleção das sequências, a afinidade de cada uma com os HLAs previamente selecionados foi testada através do NetCTLpan. Os possíveis epítomos encontrados foram classificados com base no %-rank, sendo selecionados os possíveis epítomos com %-rank ≤ 0.15 . Foram selecionados para análise neste estudo apenas os quinze possíveis epítomos que se ligam a uma maior quantidade de moléculas de HLA, sendo cinco deles para HLA-A, cinco para HLA-B e cinco para HLA-C. **RESULTADOS:** A sequência de proteína Spike da variante Delta utilizada para estudo apresentou sessenta e quatro potenciais epítomos para moléculas de HLA, sendo que, desses, sessenta e dois potenciais epítomos também estão presentes na sequência de proteína Spike de referência. Dessa forma, foi observado o surgimento de dois novos potenciais epítomos, inexistentes na sequência de proteína Spike de referência. Além disso, foram encontrados oito pontos de alto reconhecimento por HLAs nas duas sequências estudadas. **CONCLUSÃO:** Observou-se grande semelhança tanto qualitativa quanto quantitativa entre os potenciais epítomos. Apesar disso, as variações encontradas entre as sequências provavelmente possuem repercussão na maior transmissibilidade e maior gravidade do quadro clínico relatadas pela infecção pela Delta. Ademais, foi demonstrado que a infecção por diferentes variantes está provavelmente relacionada com a proteção contra reinfecção pelo SARS-CoV-2.

Palavras-chave: SARS-CoV-2. Predição. Epítomos. Linfócitos T CD8. Resposta imune.

ABSTRACT

Portela, MM. In silico prediction of sars-cov-2 delta variant cytotoxic t-lymphocyte epitopes [completion of course work]. Salvador: School of Medicine, Bahia School of Medicine and Public Health, Bahia; 2022.

INTRODUCTION: In January 2020, Chinese scientists isolated SARS-CoV-2, the virus that causes COVID-19, causing a global health crisis that has led to more than 6 million deaths. In this context, variants of the same virus emerged, especially the Delta variant, which showed an increase in the pathogenicity and transmissibility of the disease. However, to date, few studies have described the immune response triggered by cytotoxic CD8 T cells to SARS-CoV-2. Thus, it is important to identify the potential T epitopes of the Delta variant of SARS-CoV-2, since it will allow inferences about the possibility of escape from cellular immunity generated by primum infection or vaccination and what impact this escape is capable of generating on the immune system. individuals' health. **OBJECTIVE:** To predict in silico epitopes of Spike protein from CTL lymphocytes of the Delta variant of SARS-CoV-2, comparing them with epitopes of the mother sequence isolated from Wuhan. **METHODOLOGY:** For the prediction, the HLA molecules most frequently found in the Brazilian population were used, found on the website “The Allele Frequency”. The nucleotide sequences of the SARS-CoV-2 reference sequence and the SARS-CoV-2 Delta variant were obtained through the GISAID platform. The nucleotide sequence of the reference sequence was translated using the “Expase Translate” software and the amino acid sequences corresponding to the Spike protein were collected. After selecting the sequences, the affinity of each one with the previously selected HLAs was tested using NetCTLpan. The possible epitopes found were classified based on %-rank, and possible epitopes with %-rank ≤ 0.15 were selected. Only the fifteen possible epitopes that bind to a greater amount of HLA molecules were selected for analysis in this study, five of them for HLA-A, five for HLA-B and five for HLA-C. **RESULTS:** The Delta variant Spike protein sequence used for the study showed sixty-four potential epitopes for HLA molecules, and of these, sixty-two potential epitopes are also present in the reference Spike protein sequence, thus, it was observed the emergence of two new potential epitopes, lacking in the reference Spike protein sequence. In addition, eight high recognition points for HLAs were found. **CONCLUSION:** A great similarity was observed both qualitatively and quantitatively between the potential epitopes. Despite this, the variations found between the sequences probably have an impact on the greater transmissibility and greater severity of the clinical picture reported by Delta infection. Furthermore, it has been shown that infection by different variants is likely related to protection against SARS-CoV-2 reinfection.

Keywords: SARS-CoV-2; Prediction. Epitopes. CD8 T lymphocytes. Immune response.

SUMÁRIO

1. INTRODUÇÃO	6
2. OBJETIVO.....	9
3. REVISÃO DE LITERATURA.....	10
3.1. INTRODUÇÃO	10
3.2. ESTRUTURA DO SARS-COV-2	10
3.2.1 <i>ESTRUTURA DA PROTEÍNA SPIKE</i>	11
3.3. TRANSMISSÃO DO SARS-COV-2	12
3.4. INVASÃO DO SISTEMA IMUNOLÓGICO E RESPOSTA IMUNE	13
3.5. MUTAÇÕES E SURGIMENTO DE NOVAS VARIANTES	14
3.6. VARIANTE DELTA	14
3.7. IMPORTÂNCIA DE PREDIZER EPÍTOPOS TCTL DO SARS-COV-2.....	15
4. MÉTODOS.....	16
4.1. DESENHO DO ESTUDO	16
4.2. SELEÇÃO DAS MOLÉCULAS DE HLA.....	16
4.3. OBTENÇÃO DOS EPÍTOPOS	17
4.4. ANÁLISE ESTATÍSTICA	19
4.5. CRITÉRIOS DE INCLUSÃO	19
4.6. COMITÊ DE ÉTICA.....	20
6. RESULTADOS	21
6.1 POTENCIAIS EPÍTOPOS DE ALTA AFINIDADE.....	21
6.1.2. <i>ÁREAS DAS SEQUÊNCIAS COM ALTA AFINIDADE PARA MOLÉCULAS DE HLA</i>	21
6.1.2 <i>POTENCIAIS EPÍTOPOS DE ALTA AFINIDADE SELECIONADOS PARA ESTUDO</i>	22
6.2 QUANTIDADE DE EPÍTOPOS COM %-RANK ≤ 0.15 PARA MOLÉCULAS DE HLA CLASSIFICADAS COM BASE EM SUA FREQUÊNCIA NA POPULAÇÃO BRASILEIRA	25
6.3. ALINHAMENTO DAS SEQUÊNCIAS DE PROTEÍNA SPIKE ESTUDADAS.....	28
6.3.1. <i>ANÁLISE DAS MUTAÇÕES DAS SEQUÊNCIAS DE PROTEÍNA SPIKE DA VARIANTE DELTA EM RELAÇÃO À SEQUÊNCIA REFERÊNCIA</i>	28

6.4. IMPACTO DAS MUTAÇÕES NAS SEQUÊNCIAS DE PROTEÍNA SPIKE DAS VARIANTES QUANDO COMPARADAS À SEQUÊNCIA SPIKE DE REFERÊNCIA	28
6.4.1. SEQUÊNCIA DE PROTEÍNA SPIKE DE REFERÊNCIA	28
6.4.2. SEQUÊNCIA DE PROTEÍNA SPIKE DA VARIANTE DELTA	28
7. DISCUSSÃO	31
8. CONCLUSÃO.....	33
REFERÊNCIAS.....	34
APÊNDICE	37
A. POTENCIAS EPÍTOPOS ENCONTRADOS PELO NETCTLPAN PARA HLA-A, HLA-B E HLA-C DE PROTEÍNA SPIKE REFERENTE À SEQUÊNCIA DE REFERÊNCIA	37
B. POTENCIAS EPÍTOPOS ENCONTRADOS PELO NETCTLPAN PARA HLA-A, HLA-B E HLA-C NA SEQUÊNCIA DE PROTEÍNA SPIKE DA VARIANTE DELTA.....	38

1. INTRODUÇÃO

Em 31 de dezembro de 2019, as autoridades nacionais da China relataram a existência de pacientes com pneumonia de etiologia desconhecida à Organização Mundial da Saúde (OMS). Em 7 de Janeiro de 2020, cientistas chineses isolaram um vírus de pacientes em Wuhan, o qual foi identificado como um novo coronavírus causador da Síndrome Respiratória Aguda Severa (SARS).^{1,2} Assim, o vírus passou a ser conhecido como SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2), uma vez que um vírus da mesma família causou uma epidemia de SARS na China, em 2002, ficando conhecido como SARS-CoV.^{1,3,4} Em 21 de Janeiro de 2020 foi registrado o primeiro caso Da COVID-19 nos EUA e em 26 de Janeiro do mesmo, um homem testou positivo para a doença na Alemanha, demonstrando a propagação do vírus pelo América e pela Europa.⁵ Em decorrência da velocidade de contaminação, em 11 de março de 2020, a Organização Mundial da Saúde declarou surto do SARS-CoV-2 como uma pandemia global. Dados desse organização mostram que, desde o início da pandemia até 02 de setembro de 2022, 601.189.435 casos de COVID-19 haviam sido confirmados no mundo, sendo que, destes, 6.495.346 foram a óbito.⁶ No Brasil, nessa mesma data, o número de casos notificados totalizava 34.429.853, sendo 683.965 óbitos.⁶ Essa rápida e intensa propagação, associada ao alarmante número de mortes, destaca a importância de elucidar melhor o mecanismo de invasão do sistema imune e de replicação desse vírus.

O SARS-CoV-2 é um vírus de RNA de fita simples de sentido positivo⁴ e, assim como o SARS-CoV e o MERS-CoV (causador da síndrome respiratória do Oriente Médio), pertence ao gênero beta *Coronavirus* na família Coronaviridae.¹ Seu genoma codifica proteínas estruturais e não estruturais. As proteínas estruturais incluem a proteína spike (S), proteína de envelope (E), glicoproteína de membrana (M), fosfoproteína de nucleocapsídeo (N) e as proteínas não estruturais incluem 14 ORFs para replicação viral e processo de montagem.⁷ Como um vírus de RNA típico, a taxa evolutiva média para coronavírus é de aproximadamente 10^{-4} substituições de nucleotídeos por local por ano, com mutações surgindo durante cada ciclo de replicação.⁴ Apesar de maior parte dessas mutações ter pouco ou nenhum impacto sobre as características do vírus, algumas podem provocar variações importantes, como aumento da patogenicidade e escape de vacinas, dando origem às

chamadas variantes, uma vez que possuem diferenciação importante no genoma em relação à sequência-mãe. Essas variantes foram classificadas pela OMS como de interesse e de preocupação, sendo que as últimas apresentam um impacto significativo na transmissibilidade, gravidade e/ou imunidade.⁸ As variantes de preocupação descritas são a variante Alpha, detectada pela primeira vez no final de setembro de 2020, a variante Beta, detectada em setembro de 2020, a Gama, detectada em janeiro de 2021, a variante Delta, detectada em dezembro de 2020 e a Ômicron, detectada em novembro de 2021.⁹ A variante Delta, foco do presente estudo, é 60% mais transmissível que a variante alfa e tornou-se, em cerca de 8 meses, a cepa dominante do SARS-CoV-2.¹⁰

Os vírus são parasitas intracelulares obrigatórios e a resposta imune a esses microrganismos inclui componentes da imunidade inata e adaptativa. Os mecanismos da imunidade inata são a inibição por interferons do tipo I e a morte de células infectadas por linfócitos NK enquanto a imunidade adaptativa é mediada por anticorpos e por linfócitos T CD8 citotóxicos (CTL).¹¹ Os linfócitos CTL reconhecem pequenas sequências de peptídeos, originados de antígenos virais citosólicos que são degradados em proteossomos e apresentados em associação à moléculas de MHC de classe I. Ao reconhecer as células infectadas, As CTL, são capazes de destruir por apoptose qualquer célula nucleada infectada.¹¹

Sabe-se que as variantes dos SARS-CoV-2 podem ocasionar um escape da resposta imune, inclusive daquela induzida pelas vacinas, como ocorre com a variante Delta, que possui maior patogenicidade e transmissibilidade quando comparada à sequência-mãe isolada de Wuhan.¹⁰ Poucos estudos descreveram a resposta imune celular, que envolve as células T CD8 citotóxicas. Dessa forma, destaca-se a importância de identificar os epítomos T potenciais da variante Delta do SARS-CoV-2.¹⁰ Utilizando ferramentas de bioinformática é possível prever epítomos e identificar regiões imunodominantes das sequências da variante Delta, comparando-a sequência-mãe isolada de Wuhan. Pode-se igualmente observar se ocorreram mutações em regiões alvo da resposta imune mediada por linfócitos TCTL.¹²

A predição dos epítomos TCTL das variantes do SARS-CoV-2 permitirá inferir sobre a possibilidade de escape da imunidade celular gerada pela primeira infecção ou pela vacinação

e qual impacto esse escape é capaz de gerar na saúde dos indivíduos.

2. OBJETIVO

Predizer *in silico* epítomos da proteína Spike de linfócitos CTL da variante Delta do SARS-CoV-2, comparando-os com epítomos da sequência-mãe isolada de Wuhan.

3. REVISÃO DE LITERATURA

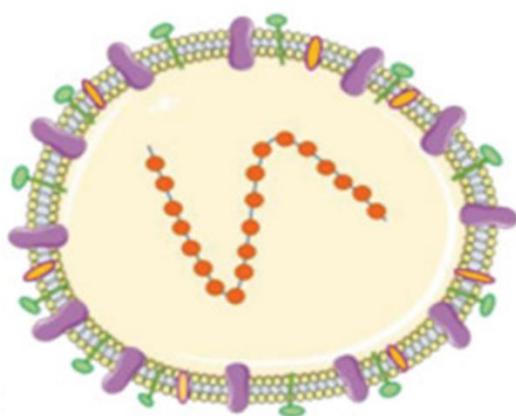
3.1. INTRODUÇÃO

No final de 2019, unidades de saúde relataram pacientes com pneumonia de causa desconhecida em Wuhan, na China.¹³ Um grupo de pesquisadores chineses rapidamente identificou o agente causador se tratava como um novo coronavírus, o qual foi nomeado como SARS-CoV-2 (Coronavírus 2 causador da Síndrome Respiratória Aguda Grave).¹³ Esse novo vírus, altamente transmissível, provocou uma crise de saúde pública global sem precedentes, preocupando autoridades de todos os países.¹⁴

3.2. ESTRUTURA DO SARS-COV-2

O SARS-CoV-2 faz parte da família dos betacoronavirus, que constitui um grupo de vírus de RNA de fita simples com sentido positivo, com aproximadamente 29,9kb.^{7,15} Seu genoma codifica 4 proteínas estruturais, a proteína do nucleocapsídeo (N), que forma o capsídeo fora do genoma, o qual é recoberto por um envelope associado às outras três proteínas estruturais, e as proteínas spike (S), proteína de membrana (M) e proteína de envelope (E).¹⁵ A estrutura do SARS-CoV-2 está representada na figura 1.

Figura 1 – Estrutura do SARS-CoV-2.



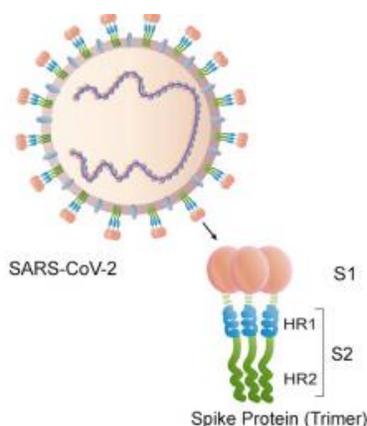
Proteínas estruturais	Desenho esquemático
Membrana	
Envelope	
Nucleocapsídeo	
Spike	
ssRNA	

Fonte: Mohamadian M, Chiti H, Shoghli A, Biglari S, Parsamanesh N, Esmailzadeh A. COVID-19: Virology, biology and novel laboratory diagnosis.

3.2.1 ESTRUTURA DA PROTEÍNA SPIKE

A proteína Spike medeia a fixação do SARS-COV-2 à membrana da célula hospedeira e a fusão durante a infecção viral.¹⁶ Ela é uma proteína trimérica formada por 1273 aminoácidos e consiste em um peptídeo sinal (aminoácidos 1-13), a subunidade S1 (resíduos 14-685), responsável pela ligação ao receptor e a subunidade S2 (resíduos 686-1273), responsável pela fusão à membrana.¹⁶ A estrutura da proteína Spike está representada na figura 2.

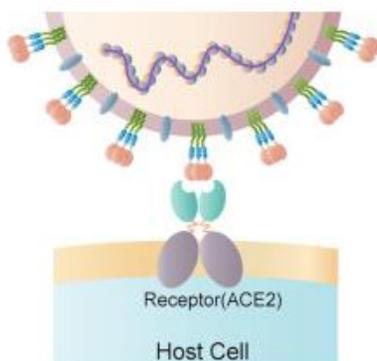
Figura 2 – Estrutura da proteína Spike.



Fonte: Huang Y, Yang C, Xu X-F, Xu W, Liu S-W. Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19.

A subunidade S1 consiste ainda em um domínio N-terminal (NTD) (resíduos 14-305) e o domínio de ligação ao receptor (RBD) (resíduos 319-541).^{16,17} O NTD é responsável pela regulação da conformação da proteína Spike e algumas mutações nesse domínio resultam no aumento da entrada de células correlacionadas com a maior apresentação de RBDs aos receptores ACE2.¹⁷ Já a subunidade S2 contém o peptídeo de fusão (FP), sequências repetidas de heptapeptídeos (HR1 e HR2), domínio transmembranar (TM) e domínio citoplasmático (CT).^{16,17} O FP é responsável por ancorar ao receptor alvo, facilitando a entrada do vírus por romper as bicamadas lipídicas das células alvo.¹⁷ HR 1 e HR2 são importantes para a fusão e entrada das subunidade S2 e o domínio TM ancora a proteína S à proteína M viral e a subunidade S2 na cauda do CT.¹⁷ A ligação da proteína S ao receptor ACE2 está representada na figura 3.

Figura 3 – Ligação da proteína S ao receptor ACE2.



Fonte: Huang Y, Yang C, Xu X-F, Xu W, Liu S-W. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19.

Essa proteína é essencial para a entrada do vírus nas células hospedeiras, fazendo desta um atraente alvo antiviral.^{15,18} Vacinas como a Pfizer/BioNTech e Moderna, aprovadas pela FDA, usam um mRNA que codifica uma proteína Spike modificada estabilizada, de forma a induzir a ação do sistema imunológico contra a proteína Spike antes de sua fusão com a membrana celular.¹⁹

3.3. TRANSMISSÃO DO SARS-COV-2

A transmissão do SARS-CoV-2, de humano para humano, ocorre através da exposição às gotículas respiratórias ou aos aerossóis provenientes de indivíduos infectados, numa taxa extremamente alta.²⁰ Dentro do corpo, o domínio de ligação ao receptor (RBD) da proteína S liga-se aos receptores da enzima conversora de angiotensina (ACE2) do hospedeiro e entra nas células, onde sofre replicação, transcrição e tradução^{7,13}. Dessa maneira, o SARS-CoV-2 utiliza de uma proteína própria do hospedeiro para infectá-lo, de forma a escapar do reconhecimento pelo sistema imune.⁷ No interior das células, o SARS-CoV-2 utiliza diversos outros mecanismos de escape para a resposta imunológica, como o silenciamento do sistema de tradução, degradação do RNA mensageiro do hospedeiro e formação de um cap viral semelhante ao do hospedeiro. Juntos, esses fatores contribuem para dificultar o reconhecimento do genoma viral pelo sistema imunológico.⁷

3.4. INVASÃO DO SISTEMA IMUNOLÓGICO E RESPOSTA IMUNE

Quando antígenos do SARS-CoV-2 são reconhecidos, desencadeiam respostas imunes inatas e adaptativas no organismo hospedeiro.⁷ A resposta imune inata é desencadeada por receptores de reconhecimento padrão (PPRs), que provocam uma cascata de sinalização que culmina na produção de citocinas, interferons, fator de necrose tumoral alfa, interleucinas 1 e 6, além de outras citocinas que irão induzir atividade antiviral e ativação da imunidade adaptativa.²¹ Já a resposta imune humoral é mediada por anticorpos direcionados às glicoproteínas de superfície viral, principalmente às proteínas S e N, entretanto ainda não está esclarecido se essa resposta fornece imunidade duradoura.²¹

Já a resposta imune mediada por células T, foco do presente estudo, foi demonstrada em pacientes convalescentes (o que evidencia uma resposta duradoura), uma vez que foram encontradas células T CD4+ e T CD8+ trabalhando em conjunto para eliminar células infectadas.²¹

Essa resposta depende da apresentação de antígenos pelas moléculas de MHC (Complexo Principal de Histocompatibilidade) para os linfócitos. O MHC consiste na região genética responsável pela produção de antígenos/proteínas encontradas na superfície das células.¹¹ Esses antígenos/proteínas são conhecidos como Antígenos Leucocitários Humanos (moléculas de HLA) e são essenciais para a apresentação de antígenos de linfócitos TCD8+.¹¹

As moléculas de HLA são codificadas por genes polimórficos.²² Sendo assim, as moléculas codificadas por um mesmo gene apresentam muitas variantes, encontradas na população em diversas frequências, o que pode explicar porque cada indivíduo responde de maneira diferente à infecção pelo mesmo patógeno.²²

Dessa forma, destaca-se a importância de considerar a prevalência de diferentes moléculas de HLA em determinada população (no presente estudo será estudada a população brasileira) quando se pretende caracterizar a resposta imune mediada por linfócitos TCD8.

3.5. MUTAÇÕES E SURGIMENTO DE NOVAS VARIANTES

Assim como outros vírus de RNA, o SARS-CoV-2 sofre mutações constantemente, pois, quanto mais chances de um vírus se expandir, mais ele sofre replicações e, conseqüentemente, mutações.^{21,10} Estima-se que o SARS-CoV-2 possa fazer de 1 bilhão de cópias após infectar um indivíduo e, durante essas replicações, ocorrem erros que podem se manifestar como mutações capazes de gerar novas variantes.¹⁰

Variantes são sequências de SARS-CoV-2 que diferem da sequência original em decorrência de mutações genéticas.²³ Em dezembro de 2020, a Organização Mundial da Saúde, caracterizou essas variantes de em variantes de interesse (VOIs) e variantes de preocupação (VOCs), sendo as VOIs sequências que sofreram alterações que podem afetar a implicabilidade ou a transmissão da doença.¹⁰ Elas passam a ser VOCs caso haja um aumento na transmissibilidade do vírus ou mudança prejudicial na epidemiologia do COVID-19, aumento virulência ou alteração na apresentação clínica da doença ou diminuição da eficácia nas medidas de saúde pública e sociais ou diagnósticos, vacinas ou tratamentos.¹⁰

3.6. VARIANTE DELTA

A variante Delta, que consiste no foco do presente estudo, foi relatada pela primeira vez na Índia, em dezembro de 2020 e é classificada pela OMS como uma VOC.¹⁰ Essa variante é de 60% mais transmissível que a variante alfa e tornou-se, em cerca de 8 meses, a cepa dominante do SARS-CoV-2.²⁴ Sua alta transmissibilidade, deve-se, provavelmente, à sua maior capacidade de invadir o sistema imune do indivíduo infectado quando comparada à cepa original.²⁴ Suspeita-se que a grande capacidade de invasão dessa cepa decorra de mutações encontradas na proteína spike, as quais podem induzir mudanças bioquímicas e na estrutura do vírus, facilitando a fusão da proteína com a célula hospedeira ou inibindo seu reconhecimento por anticorpos.¹⁰ Uma vez que o alvo das principais vacinas desenvolvidas contra a COVID-19 é a proteína Spike, mutações nessa região do genoma podem comprometer a eficácia dessas vacinas.²³

3.7. IMPORTÂNCIA DE PREDIZER EPÍTOPOS TCTL DO SARS-COV-2

A relevância de predizer epítomos TCTL da variante Delta e compará-los com epítomos T da sequência original é a capacidade de prever escape da resposta imune, inclusive daquela induzida por vacinas.²⁵ Uma vez que sejam preditos tais epítomos, será possível avaliar o impacto das mutações na resposta mediada por linfócitos TCTL e de que forma isso influencia no curso e na gravidade da doença causada pelo SARS-CoV-2 ^{21,25}.

4. MÉTODOS

4.1. DESENHO DO ESTUDO

Trata-se de um estudo analítico e descritivo, que visa prever epítomos de linfócitos T da variante Delta e compará-los aos epítomos da sequência-mãe do SARS-CoV-2.

4.2. SELEÇÃO DAS MOLÉCULAS DE HLA

As moléculas de HLA com maior frequência na população brasileira foram utilizadas para prever os epítomos de linfócitos T da sequência-mãe e da variante Delta do SARS-CoV-2. Essas moléculas foram encontradas através do site “The Allele Frequency”. Foi realizada uma busca das moléculas de HLA presentes na população brasileira e da frequência que aparecem. Os HLAs encontrados e utilizados no estudo estão representados na tabela 1.

Quadro 1 - HLAs utilizados no estudo

HLA A	HLA B	HLA C
A*0201	B*3501	Cw-0401
A*0301	B*4403	Cw-0701
A*2301	B*5101	Cw-1601
A*0101	B*0702	Cw-0602
A*3001	B*5301	Cw-0304
A*1101	B*1402	Cw-0802
A*2402	B*4402	Cw-0702
A*6801	B*5201	Cw-0501
A*3402	B*1801	Cw-1203
A*3601	B*0801	Cw-0303
A*6802	B*5801	Cw-0202
A*7401	B*4201	Cw-1701
A*3002	B*1501	Cw-0210
A*2601	B*1510	Cw-1801

Quadro 1 - HLAs utilizados no estudo

A*3301	B*4901	Cw-1402
A*0202	B*1503	Cw-1502
A*2902	B*4001	Cw-0102
A*3201	B*8101	Cw-1202
A*3101	B*5703	Cw-1505
A*0205	B*4501	Cw-1602
A*6602	B*3801	Cw-0704
A*2901	B*4002	Cw-0804
A*8001	B*5501	Cw-0103
A*3004	B*5701	Cw-0302
A*0302	B*5802	Cw-0705
A*2403	B*1302	Cw-0707
A*2501	B*1401	Cw-0813
A*3303	B*1516	Cw-1704
A*0222	B*2705	
A*0234	B*3503	
A*6603	B*3508	
A*7403	B*4004	
A*6602	B*5001	
A*0102	B*0705	
A*0211	B*1517	
A*6901	B*3504	
A*7402	B*3701	

Fonte: (<http://www.allelefreqencies.net/>)

4.3. OBTENÇÃO DOS EPÍTOPOS

As sequências de nucleotídeos da sequência-referência do SARS-CoV-2 e da variante Delta do SARS-CoV-2 foram obtidas através da plataforma GISAID, uma plataforma que promove o compartilhamento rápido de dados de todos os vírus da gripe e do coronarírus que causam a COVID-19. Na página online do GISAID, encontra-se a aba “EpiCoV”, onde é possível pesquisar entre as sequências registradas.

A sequência de nucleotídeos da sequência-referência foi traduzida com auxílio do software “Expase Translate” e foram coletadas as sequências de aminoácidos correspondentes à proteína Spike.

Foi selecionada uma sequência da proteína Spike utilizada como referência e uma sequência de proteína Spike para a variante Delta. A sequência referência utilizada foi isolada pelo Instituto de Virologia de Wuhan em 30 dezembro de 2019, a partir de uma amostra de lavado broncoalveolar coletado no Hospital Wuhan Jinyintan. A sequência Spike Delta (ID de acesso EPI_ISL_6661895) foi coletada em 12/03/2020 em Nova York, Estados Unidos.

Além da sequência de nucleotídeos, o GISAID demonstra as mutações (substituições, deleções e inserções de aminoácidos) que ocorreram em cada sequência em relação à sequência de proteína Spike utilizada como referência. Essas mutações foram alinhadas e comparadas com a sequência referência a fim de gerar as sequências de aminoácidos correspondentes às sequências selecionadas como amostras.

Após a seleção das sequências, a afinidade de cada uma com os HLAs previamente selecionados foi testada através do NetCTLpan, um software que prevê epítomos CTL em sequências de proteínas. As previsões podem ser feitas para peptídeos com comprimento de oito a onze resíduos de aminoácidos, sendo que a maioria das moléculas de HLA tem uma forte preferência pela ligação com epítomos de nove resíduos aminoácidos. O método de previsão do NetCTLpan integra a previsão da ligação do MHC de classe I, clivagem proteossomal C terminal e eficiência de transporte TAP.²⁶ Os valores de previsão são calculados como uma média ponderada das pontuações de clivagem do terminal MHC, TAP e C. e como %-rank para um conjunto de duzentos mil peptídeos naturais aleatórios.²⁶

Após realizar a predição no NetCTLpan, os possíveis epítomos foram classificados com base no %-rank, que consiste na classificação da afinidade prevista de determinado HLA em comparação com um conjunto de peptídeos naturais aleatórios, ou seja, é um valor numérico que representa a força da ligação entre o possível epítomo e determinado HLA.²⁶ Quanto menor for esse valor, maior será a força de ligação. Os ligantes fortes são definidos como tendo %-rank < 0.5 e os ligantes fracos com %-rank < 2.²⁶ Para este estudo foram selecionados os possíveis epítomos com %rank \leq 0.15, pois, na prática, ao se utilizar

peptídeos para estímulo de células *in vitro*, os peptídeos que tem melhores resultados são aqueles cujo %-rank é menor ou igual a 0.15.

Dentre os epítomos previstos com %-rank ≤ 0.15 pelo NetCTLpan, foram selecionados para análise neste estudo apenas os quinze possíveis epítomos que se ligam a uma maior quantidade de moléculas de HLA, sendo cinco deles para HLA-A, cinco para HLA-B e cinco para HLA-C.

Os possíveis epítomos escolhidos foram os mesmos para as duas sequências, entretanto sua posição em cada sequência pode sofrer variações conforme as mutações intrínsecas às mesmas.

4.4. ANÁLISE ESTATÍSTICA

Foi utilizado o valor das pontuações combinadas (comb) para comparar os quinze possíveis epítomos selecionados. Esse valor consiste na combinação entre o processamento do possível epítomo, sua apresentação e o transporte pelo TAB, etapas necessárias para que a resposta imune ocorra.

Os cinco possíveis epítomos para HLA, os cinco possíveis epítomos para HLA-B e os cinco possíveis epítomos para HLA-C foram comparados separadamente através da média aritmética simples, em decorrência da pequena amostragem. Seu cálculo foi realizado através do GraphPad, utilizando o teste ANOVA Kruskal Wallis seguido do pós-teste de Dunn's. Essa análise foi empregada para avaliar se as mutações provocaram alguma alteração no processamento e apresentação do peptídeo.

4.5. CRITÉRIOS DE INCLUSÃO

Foram incluídos no estudo os potenciais epítomos previstos pelo NetCTLpan com %-rank ≤ 0.15 .

4.6. *COMITÊ DE ÉTICA*

Por se tratar de um estudo descritivo, esse estudo não precisou ser submetido ao Comitê de Ética em Pesquisa (CEP).

6. RESULTADOS

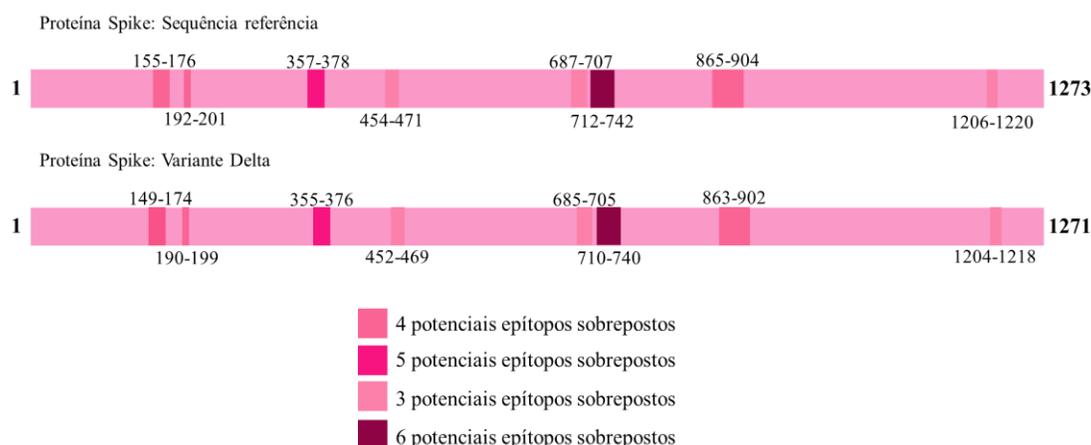
6.1 POTENCIAIS EPÍTOPOS DE ALTA AFINIDADE

Inicialmente foram selecionados potenciais epítomos que provavelmente serão apresentados a certas moléculas de HLA A, HLA B e HLA C previamente selecionadas (tabela 1) que são encontradas na população brasileira. Foram escolhidos os potenciais epítomos com %-rank menor ou igual a 0.15 para as duas sequências de proteína Spike estudadas. Para a sequência de proteína Spike utilizada como referência, foram encontrados 33 potenciais epítomos para HLA A, 27 potenciais epítomos para HLA B e 21 potenciais epítomos para HLA C (Anexo A). Já para a sequência de proteína Spike da variante Delta (EPI_ISL_6661895) foram encontrados 34 potenciais epítomos para HLA A, 27 potenciais epítomos para HLA B e 19 potenciais epítomos para HLA C (Anexo A).

6.1.2. ÁREAS DAS SEQUÊNCIAS COM ALTA AFINIDADE PARA MOLÉCULAS DE HLA

A partir do alinhamento das sequências, foram encontrados oito pontos com mais de três peptídeos sobrepostos ou com mais de três peptídeos na mesma região em ambas as variantes, esses pontos foram considerados como pontos de alto reconhecimento por HLAs. Esses pontos ocupam as seguintes posições na sequência referência: (1) 155 a 176, com quatro potenciais epítomos sobrepostos; (2) 192 a 201, com 4 potenciais epítomos sobrepostos; (3) 357 a 378, com cinco potenciais epítomos sobrepostos; (4) 454 a 471, com três potenciais epítomos sobrepostos; (5) 687 a 707, com três potenciais epítomos sobrepostos; (6) 712 a 742, com seis potenciais epítomos sobrepostos; (7) 865 a 904, com quatro potenciais epítomos sobrepostos; e (8) 1206 a 1220, com três potenciais epítomos sobrepostos.

Na variante Delta também foram encontrados oito pontos com mais de três peptídeos sobrepostos, entretanto, num dos pontos de alto reconhecimento, ocorreram duas deleções (F157delA e R158del). Essas deleções provocaram o reconhecimento de dois novos peptídeos (posições 149-158 e 152-161 na sequência Delta) e fizeram com que um peptídeo deixasse de ser reconhecido (posição 155-164). A representação gráfica da proteína Spike das duas sequências e a localizações que os pontos de alto reconhecimento ocupam na sequência estão representadas na figura 4.

Figura 4 – Representação gráfica da proteína Spike e dos pontos de alto reconhecimento.

Fonte: Arquivo do autor.

6.1.2 POTENCIAIS EPÍTOPOS DE ALTA AFINIDADE SELECIONADOS PARA ESTUDO

Dentre os potenciais epítopos encontrados para as duas sequências, foram escolhidos os epítopos com maior quantidade de ligações com $\% \text{-rank} \leq 0.15$ que se repetiam nas duas sequências de proteína Spike estudadas. Sua posição no interior das sequências variou em função das mutações que cada uma sofreu, entretanto, a sequência de resíduos de aminoácidos foi conservada.

Para sequência a referente à proteína Spike da sequência referência, os potenciais epítopos selecionados ocupam as posições: (1) 360, 161, 257, 301 e 453 para HLA-A; (2) 868, 895, 83, 159 e 235 para HLA-B; e (3) 868, 885, 1, 159 e 368 para HLA-C. Para a sequência referente à proteína Spike da variante Delta, os potenciais epítopos selecionados ocupam as posições: (1) 358, 159, 255, 299 e 451 para HLA-A; (2) 866, 893, 83, 157 e 233 para HLA-B; e (3) 866, 883, 1, 157 e 366 para HLA-C.

As posições dos possíveis epítopos que se ligam ao maior número de HLAs com alta afinidade ($\% \text{-rank} \leq 0.15$) para as duas sequências de proteína Spike (referência e Delta) selecionadas para estudo estão representados no quadro 2.

Quadro 2 - Potenciais epítopos selecionados para estudo e quantidades de ligações de alta afinidade que cada um realiza para HLA-A, HLA-B e HLA-C (continua).

Sequência referência – Spike					
HLA-A		HLA-B		HLA-C	
Posição na sequência	Quantidade de ligações de alta afinidade	Posição na sequência	Quantidade de ligações de alta afinidade	Posição na sequência	Quantidade de ligações de alta afinidade
360	10	868	7	868	15
161	7	895	6	885	9
257	7	83	5	1	7

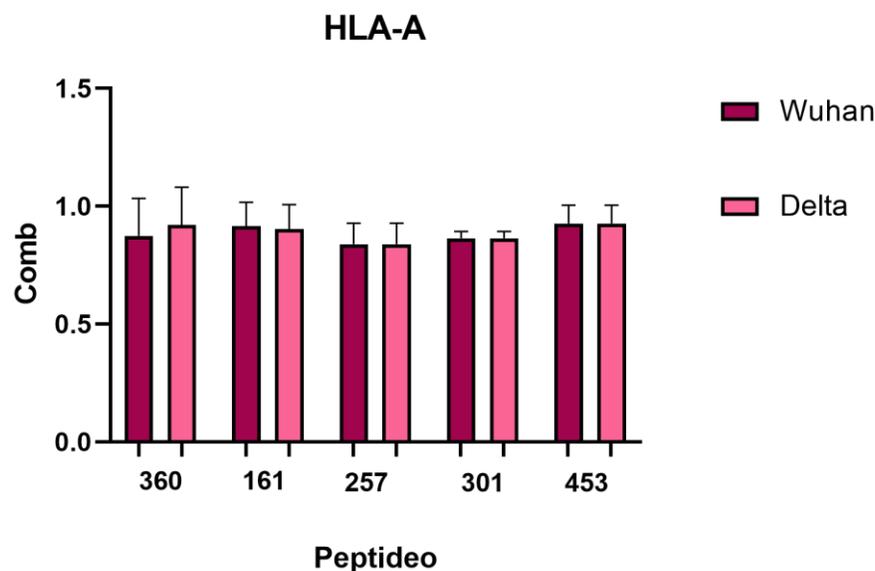
Quadro 2 - Potenciais epítomos selecionados para estudo e quantidades de ligações de alta afinidade que cada um realiza para HLA-A, HLA-B e HLA-C.

301	6	159	4	159	7
453	6	235	4	368	4
EPI_ISL_6661895 – Delta – Spike					
HLA-A		HLA-B		HLA-C	
Posição na sequência	Quantidade de ligações de alta afinidade	Posição na sequência	Quantidade de ligações de alta afinidade	Posição na sequência	Quantidade de ligações de alta afinidade
358	10	866	7	866	15
159	7	893	6	883	9
255	7	83	5	1	7
299	6	157	4	157	7
451	6	233	4	366	4

Fonte: Arquivo do autor

Os valores das médias das pontuações combinadas (comb) para os potenciais epítomos das duas sequências de proteína Spike estudadas sofreram pouca ou nenhuma variação entre as diferentes sequências, ou seja, foi observado um alto grau de conservação. Para HLA-A as médias dos valores de comb encontradas estão entre 0.5 e 1. O gráfico que compara as médias das pontuações combinadas (comb) dos potenciais epítomos para HLA-A das duas sequências de proteína estudadas está representado na figura 5.

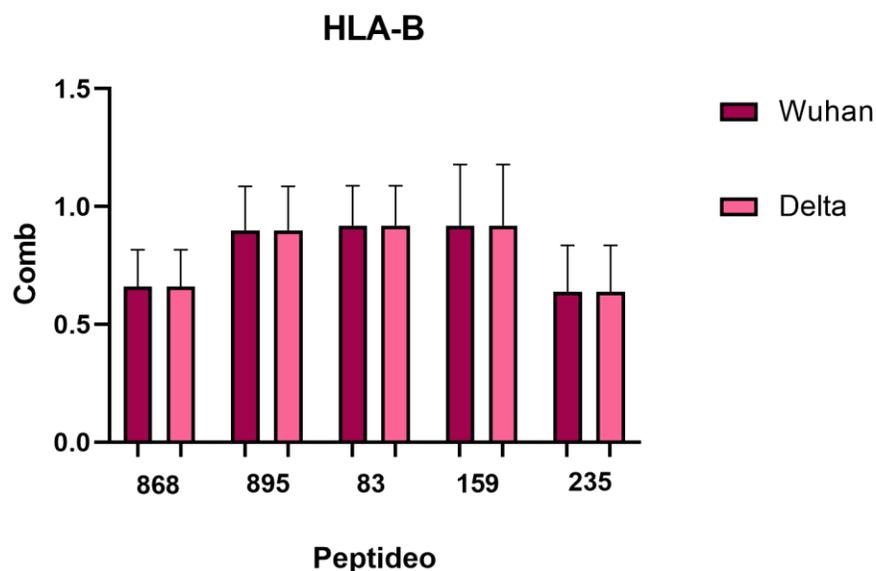
Figura 5 - Médias das pontuações combinadas (comb) dos potenciais epítomos para HLA-A.



Fonte: Arquivo do autor.

Para HLA-B, as médias dos valores de comb encontradas também variam de 0.5 a 1. O gráfico que compara as médias das pontuações combinadas (comb) dos potenciais epítomos para HLA-B das duas sequências de proteína estudadas está representado na figura 6.

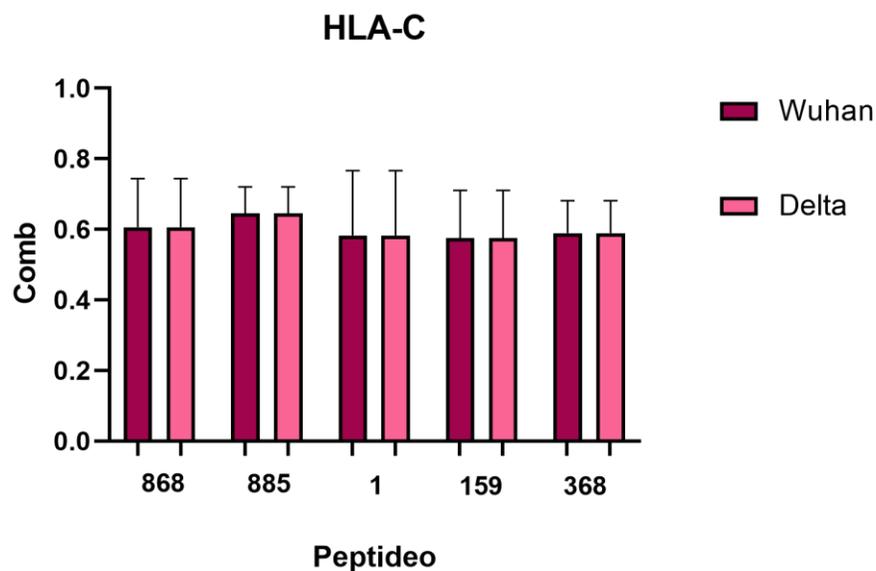
Figura 6 - Médias das pontuações combinadas (comb) dos potenciais epítomos para HLA-B.



Fonte: Arquivo do autor.

Já para HLA-C, as médias C estão entre 0.5 e 0.8. O gráfico que compara as médias das pontuações combinadas (comb) dos potenciais epítomos para HLA-C das duas sequências de proteína estudadas está representado na figura 7.

Figura 7 - Médias das pontuações combinadas (comb) dos potenciais epítomos para HLA-C.



Fonte: Arquivo do autor.

6.2 QUANTIDADE DE EPÍTOPOS COM %-RANK ≤ 0.15 PARA MOLÉCULAS DE HLA CLASSIFICADAS COM BASE EM SUA FREQUÊNCIA NA POPULAÇÃO BRASILEIRA

A quantidade de potenciais epítomos com alta afinidade (%-rank ≤ 0.15) para cada molécula de HLA foi comparada à sua frequência na população brasileira.

Quando analisadas as moléculas de HLA-A, observou-se que o HLA-A que possui maior frequência na população brasileira (HLA-A 02:01) possui pouca quantidade de ligantes com alta afinidade, sendo apenas um ligante com %-rank ≤ 0.15 para a sequência de proteína Spike da referência e nenhum ligante para a sequência de proteína Spike da variante Delta. Nota-se também uma variação na quantidade de ligantes de alta afinidade entre as duas sequências estudadas. Esse padrão se repete para o segundo HLA-A mais frequente na população brasileira (HLA-A 03:01), para o qual foram encontrados apenas três potenciais epítomos de alta afinidade para a sequência de proteína Spike da sequência referência e três para a sequência de proteína Spike da variante. Além disso, observou-se que o HLA-A 34:02 foi o que se ligou a um maior número de peptídeos, sendo sete para a sequência de proteína Spike de referência e sete para a Delta, como pode ser observado na Tabela 1.

Tabela 1 - % de HLA-A na população brasileira e número com %-rank ≤ 0.15 (continua).

HLA A	% na população	Referência	Delta
A*0201	20,7	1	0
A*0301	9,7	3	3
A*2301	6,6	1	1
A*0101	5,8	5	5
A*3001	5,1	1	1
A*1101	4,4	2	2
A*2402	4,4	2	2
A*6801	3,2	4	4
A*3402	3,1	7	7
A*3601	3,1	7	7
A*6802	2,9	3	3
A*7401	2,7	4	4
A*3002	2,6	4	4
A*2601	2,4	3	3
A*3301	2,0	0	0
A*0202	1,9	3	3
A*2902	1,7	3	3
A*3201	1,7	2	2
A*3101	1,5	2	2
A*0205	1,4	5	5
A*6602	1,2	3	3
A*2901	1,0	2	2
A*8001	0,9	1	1
A*3004	0,7	3	3
A*0302	0,5	4	4

Fonte: Arquivo do autor.

Tabela 1 - % de HLA-A na população brasileira e número de ligantes com %-rank ≤ 0.15 .

HLA A	% na população	Referência	Delta
A*2403	0,5	2	2
A*2501	0,5	2	2
A*3303	0,5	1	1
A*0222	0,3	2	2
A*0234	0,3	2	2
A*6603	0,3	4	4
A*7403	0,3	4	4
A*6602	0,2	3	3
A*0102	0,2	7	7
A*0211	0,2	1	1
A*6901	0,2	2	2
A*7402	0,2	4	4

Fonte: Arquivo do autor.

Já para as moléculas de HLA-B, observou-se que o HLA-B que possui maior frequência na população brasileira (HLA-B 35:01) também possui o maior número de ligantes com %-rank ≤ 0.15 (seis ligantes para sequência de referência e seis para Delta), como pode ser observado na tabela 2.

Tabela 2 - % de HLA-B na população brasileira e ligantes com %-rank ≤ 0.15 (continua).

HLA B	% na população	Referência	Delta
B*3501	9	6	6
B*4403	7,2	1	1
B*5101	6,6	3	3
B*0702	6,3	0	0
B*5301	5,6	4	4
B*1402	4,7	2	2
B*4402	4,3	0	0
B*5201	3,8	2	2
B*1801	3,6	0	0
B*0801	3	2	2
B*5801	3	5	5
B*4201	2,9	2	2
B*1501	2,7	0	0
B*1510	2,5	2	2
B*4901	2,5	1	1
B*1503	2	1	1
B*4001	2	2	2
B*8101	2	2	2
B*5703	1,8	1	1
B*4501	1,4	2	2
B*3801	1,3	1	1
B*4002	1,3	1	1
B*5501	1,1	3	3

Fonte: Arquivo do autor.

Tabela 2 - % de HLA-B na população brasileira e número de ligantes com %-rank ≤ 0.15 .

HLA B	% na população	Referência	Delta
B*5701	1,1	2	2
B*5802	1,1	3	3
B*1302	0,9	1	1
B*1401	0,9	2	2
B*1516	0,9	2	2
B*2705	0,9	1	1
B*3503	0,9	4	4
B*3508	0,9	2	2
B*4004	0,9	1	1
B*5001	0,9	2	2
B*0705	0,7	0	0
B*1517	0,7	3	3
B*3504	0,7	3	3
B*3701	0,7	0	0

Fonte: Arquivo do autor.

Por fim, para as moléculas de HLA-C, observou-se o que possui maior frequência na população brasileira (HLA-C 04:01) possui pouca quantidade de ligantes com alta afinidade, sendo um ligante de alta afinidade para a sequência de proteína Spike de referência e um para a Delta. Além disso, observou-se que o HLA-C com maior número de ligantes foi o HLA-C 01:02, com oito ligantes com alta afinidade para cada uma das sequências estudadas, como pode ser observado na tabela 3.

Tabela 3 - % de HLA-C na população brasileira e ligantes com %-rank ≤ 0.15 (continua).

HLA C	% na população	Referência	Delta
Cw-0401	22,2	1	1
Cw-0701	11,5	2	2
Cw-1601	7,5	5	5
Cw-0602	6,8	0	0
Cw-0304	6,6	3	3
Cw-0802	5,6	2	2
Cw-0702	5	1	1
Cw-0501	4,7	0	0
Cw-1203	3,9	3	3
Cw-0303	3,8	3	3
Cw-0202	3,6	5	5
Cw-1701	3,2	2	2
Cw-0210	2,3	5	5
Cw-1801	2,3	0	0
Cw-1402	2	3	3
Cw-1502	2	3	3
Cw-0102	1,3	8	8
Cw-1202	1,1	3	3
Cw-1505	1,1	2	2

Fonte: Arquivo do autor.

Tabela 3 - % de HLA-C na população brasileira e número de ligantes com %-rank ≤ 0.15 .

HLA C	% na população	Referência	Delta
Cw-1602	0,9	2	2
Cw-0704	0,7	1	1
Cw-0804	0,7	4	4
Cw-0103	0,2	4	4
Cw-0302	0,2	3	3
Cw-0705	0,2	1	1
Cw-0707	0,2	1	1
Cw-0813	0,2	4	4
Cw-1704	0,2	2	2
Cw-0401	22,2	1	1
Cw-0701	11,5	2	2
Cw-1601	7,5	5	5
Cw-0602	6,8	0	0
Cw-0304	6,6	3	3
Cw-0802	5,6	2	2
Cw-0702	5	1	1
Cw-0501	4,7	0	0

Fonte: Arquivo do autor.

6.3. ALINHAMENTO DAS SEQUÊNCIAS DE PROTEÍNA SPIKE ESTUDADAS

6.3.1. ANÁLISE DAS MUTAÇÕES DAS SEQUÊNCIAS DE PROTEÍNA SPIKE DA VARIANTE DELTA EM RELAÇÃO À SEQUÊNCIA REFERÊNCIA

As sequências de proteína Spike utilizadas no estado foram alinhadas e observou-se que em relação à sequência de proteína Spike utilizada como referência, na sequência de proteína Spike da variante Delta, ocorreram doze mutações, sendo duas deleções, nove substituições e uma inserção.

6.4. IMPACTO DAS MUTAÇÕES NAS SEQUÊNCIAS DE PROTEÍNA SPIKE DAS VARIANTES QUANDO COMPARADAS À SEQUÊNCIA SPIKE DE REFERÊNCIA

6.4.1. SEQUÊNCIA DE PROTEÍNA SPIKE DE REFERÊNCIA

Quando analisada a sequência de proteína Spike de referência utilizada para estudo, observou-se a presença de 66 potenciais epítomos para moléculas de HLA.

6.4.2. SEQUÊNCIA DE PROTEÍNA SPIKE DA VARIANTE DELTA

A sequência de proteína Spike da variante Delta utilizada para estudo apresentou sessenta e quatro potenciais epítomos para moléculas de HLA, sendo que, desses, sessenta e dois potenciais epítomos também estão presentes na sequência de proteína Spike de referência. Em relação à sequência de proteína Spike de referência, nenhum potencial epítomo sofreu mutação mas continua sendo reconhecido e dois potenciais epítomos são exclusivos da variante Delta. Além disso, três sequências de resíduos de aminoácidos não são reconhecidas como potenciais epítomos. Esses valores estão representados no quadro 6.

Quadro 6 - Quantidade de potenciais epítomos da sequência de proteína Spike da variante Delta.

Sequência referência	VARIANTE DELTA				
Potenciais epítomos	Potenciais epítomos	Potenciais epítomos em comum	Potenciais epítomos que sofreram mutação*	Potenciais epítomos exclusivos**	Sequências de resíduos de aminoácidos não reconhecidos***
Sessenta e seis	Sessenta e quatro	Sessenta e dois	Zero	Dois	Três

Fonte: Arquivo do autor. *Epítomos que sofreram mutação, mas continuam sendo reconhecidos. **Epítomos exclusivos reconhecidos a partir de mutações na respectiva variante. ***Sequências de resíduos de aminoácidos que pararam de ser reconhecidos devido a mutações ou que não são observados na variante.

Três mutações intrínsecas à sequência de proteína Spike da variante Delta estudada (E156G, F157del e R158del) implicaram no surgimento de dois novos potenciais epítomos, inexistentes na sequência de proteína Spike de referência. Os novos epítomos ocupam as posições 149-158 e 152-161 e observou-se que realizam ligações de alta afinidade com moléculas de HLA prevalentes na população brasileira. O peptídeo 149-158 liga-se com alta afinidade com 8 moléculas de HLA prevalentes na população brasileira, enquanto o peptídeo 152-161 liga-se a 2 moléculas de HLA com alta afinidade. Os novos potenciais epítomos, as mutações que provocaram seu reconhecimento e sua relação com os HLAs estão representadas no quadro 7.

Quadro 7 - Novos epítomos da sequência Spike da variante Delta inexistentes na sequência Spike de referência.

Potenciais epítomos novos da sequência de proteína Spike da variante Delta				
Posição	Mutações	HLA-A	HLA-B	HLA-C
149	E156G F157del R158del	Ligação com alta afinidade a 4 moléculas (30:02, 30:04, 01:02, 36:01)	Ligação com alta afinidade a 4 moléculas (15:17, 15:16, 57:01, 58:02)	Nenhuma ligação com alta afinidade
152	E156G F157del R158del	Não reconhecido	Ligação com alta afinidade a 2 moléculas (45:01, 50:01)	Não reconhecido

Fonte: Arquivo do autor. *A primeira letra antes do número representa o resíduo de aminoácido na sequência referência e a letra após o número representa o resíduo de aminoácido na variante Delta.

**del: deleção do resíduo de aminoácido representado pela letra anterior à sua posição.

Por fim, notou-se ainda que o potencial epítomo de posição 154 deixou de ser reconhecido pelas moléculas de HLA em decorrência da substituição do nucleotídeo E para

o nucleotídeo G na posição 156 e das deleções dos nucleotídeos F e R nas posições 157 e 158, respectivamente. Além disso, os potenciais epítomos das posições 193 e 196, que são reconhecidos pelas moléculas de HLA para a sequência de proteína Spike de referência não são reconhecidos pelas moléculas de HLA para a variante Delta, entretanto, ambos não sofreram mutações.

7. DISCUSSÃO

O presente estudo tem como objetivo identificar os potenciais epítomos de proteína Spike de linfócitos T citotóxicos da variante Delta do SARS-CoV-2 e compará-los com os potenciais epítomos de proteína Spike de linfócitos T citotóxicos da sequência-referência do SARS-CoV-2, isolada de Wuhan em 30 de dezembro de 2019. Ao identificar os epítomos de linfócitos TCTL, foi possível perceber que houve grande conservação entre as duas variantes estudadas. A quantidade de epítomos encontrados na variante Delta foi semelhante à encontrado na sequência de referência. Além disso, o alinhamento das duas sequências permitiu visualizar que os pontos de alto reconhecimento de epítomos são semelhantes entre ambas. Por fim, foi calculada a média das pontuações combinadas (comb) das duas sequências e notou-se que seus valores foram próximos para as duas sequências estudadas.

Após a identificação dos epítomos de proteína Spike de linfócitos T das duas sequências, observou-se grande semelhança, tanto quantitativa, quanto qualitativamente. Apesar disso, para a sequência de proteína Spike da variante Delta, um pequeno número de novos epítomos passou a ser reconhecido por moléculas de HLA utilizadas no estudo, o que significa um surgimento de áreas da sequência capazes de gerar resposta imune que não existiam na sequência de referência. Sugere-se então que esse fenômeno pode ser uma das causas do aumento da patogenicidade da variante Delta. Ademais, um pequeno número de sequências de resíduos de aminoácidos antes reconhecidas como epítomos para a sequência de proteína Spike de referência deixaram de ser reconhecidos na variante Delta em decorrência das mutações. Isso faz com que tais resíduos deixem de induzir resposta imune nos indivíduos portadores das moléculas de HLA responsáveis pelo reconhecimento desses epítomos. De acordo com Prabin Baral et al²⁷, mutações na proteína Spike da variante Delta podem reduzir as afinidades de ligação para anticorpos neutralizantes. Dessa forma, é possível que essas mutações também causem alterações na resposta imune mediada por células T, o que sugere a maior gravidade das infecções pela variante Delta quando comparadas às infecções causadas pela sequência referência do SARS-CoV-2.

Ao alinhar as sequências de proteína Spike da sequência de referência e da variante Delta foi possível notar grande semelhança entre as regiões imunodominantes, que consistem em pontos de alto reconhecimento pelas moléculas de HLA utilizadas para a predição. Apesar da grande semelhança foram encontrados dois novos peptídeos que realizam ligações de alta afinidade com moléculas de HLA prevalentes na população brasileira. Sendo assim, apesar de não existirem muitas variações entre as sequências estudadas, esse fenômeno representa uma diferença importante entre elas, provavelmente com repercussão também importante na resposta imune à variante Delta, uma vez que se localizam na região de proteína Spike correspondente do domínio N-terminal (NTD) da subunidade S1, responsável pela regulação da conformação da proteína Spike.¹⁷ Sabendo que a variação da conformação da proteína Spike é essencial para a apresentação do domínio de ligação ao receptor (RBD) da Spike, mutações no NTD podem interferir nessa apresentação e facilitar a interação do vírus com o ACE2 das células hospedeiras e conseqüentemente sua entrada nessas células. Percebe-se então que essas mutações podem contribuir para o aumento da transmissibilidade (aumento

de 1,4 a 1,64 vezes quando comparado à variante Alfa), como demonstrado no estudo realizado por Choi et al.⁹ Além disso, Choi et al evidenciaram que a infecção pela variante Delta apresentou maior gravidade, estando mais associada a maiores chances de necessidade de oxigênio, admissão na UTI e morte.⁹

Sabendo que o valor do comb consiste num número que representa a combinação entre o processamento do possível epítipo, sua apresentação e o transporte pelo TAB, é esperado que a presença de um valor semelhante entre as duas variantes represente uma resposta imune semelhante entre elas. Sendo assim, o alto grau de conservação observado entre esses valores pode ser capaz de justificar a proteção contra reinfecção por SARS-CoV-2. Essa hipótese foi demonstrada em um estudo realizado por Kojima et al em 6 países diferentes, que encontrou uma redução de 90,4% do risco de reinfecção por COVID-19 por até 10 meses.²⁸ Além disso, esse estudo demonstrou que pessoas com COVID-19 anterior tiveram um nível de proteção semelhante e durável quando comparadas às pessoas vacinadas contra SARS-CoV-2.

8. CONCLUSÃO

Ao comparar os epítomos de proteína Spike da variante Delta com os epítomos de proteína Spike da sequência isolada de Wuhan foi possível perceber grande semelhança tanto qualitativa quanto quantitativa. Apesar disso, as variações encontradas entre as sequências provavelmente possuem repercussão na maior transmissibilidade e maior gravidade do quadro clínico relatadas pela infecção pela Delta. Ademais, foi demonstrado que a infecção por diferentes variantes está provavelmente relacionada com a proteção contra reinfecção pelo SARS-CoV-2.

REFERÊNCIAS

1. Ni L, Ye F, Cheng ML, Feng Y, Deng YQ, Zhao H, et al. Detection of SARS-CoV-2-Specific Humoral and Cellular Immunity in COVID-19 Convalescent Individuals. *Immunity* [Internet]. 2020;52(6):971-977.e3. Available from: <https://doi.org/10.1016/j.immuni.2020.04.023>
2. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet*. 2020;395(10223):470–3.
3. Hasöksüz M, Kiliç S, Saraç F. Coronaviruses and sars-cov-2. *Turkish J Med Sci*. 2020;50(SI-1):549–56.
4. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* [Internet]. 2020;395(10224):565–74. Available from: [http://dx.doi.org/10.1016/S0140-6736\(20\)30251-8](http://dx.doi.org/10.1016/S0140-6736(20)30251-8)
5. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. *Acta Biomed*. 2020;91(1):157–60.
6. Painel da OMS sobre o coronavírus (COVID-19) [Internet]. Available from: <https://covid19.who.int/>
7. Mohamadian M, Chiti H, Shoghli A, Biglari S, Parsamanesh N, Esmaeilzadeh A. COVID-19: Virology, biology and novel laboratory diagnosis. *J Gene Med*. 2021;23(2):1–11.
8. Organização Mundial da Saúde. Tracking SARS-CoV-2 variants [Internet]. Tracking SARS-CoV-2 variants. 2021 [cited 2021 Jun 22]. Available from: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>
9. Choi JY, Smith DM. SARS-CoV-2 variants of concern. *Yonsei Med J*. 2021;62(11):961–8.
10. Hendaus MA, Jomha FA. Delta variant of COVID-19: A simple explanation. *Qatar Med J* [Internet]. 2021;2021(3):49. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/34660217>
11. Abbas AK. *Imunologia Celular e Molecular*. 9th ed. Grupo GEN, editor. 2019.
12. Kiyotani K, Toyoshima Y, Nemoto K, Nakamura Y. Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2. *J Hum Genet* [Internet]. 2020;65(7):569–75. Available from: <http://dx.doi.org/10.1038/s10038-020-0771-5>
13. Hu B, Guo H, Zhou P, Shi ZL. Characteristics of SARS-CoV-2 and COVID-19. *Nat Rev Microbiol* [Internet]. 2021;19(3):141–54. Available from: <http://dx.doi.org/10.1038/s41579-020-00459-7>
14. Gustine JN, Jones D. Immunopathology of Hyperinflammation in COVID-19. *Am J Pathol*. 2021;191(1):4–17.

15. Wang MY, Zhao R, Gao LJ, Gao XF, Wang DP, Cao JM. SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. *Front Cell Infect Microbiol.* 2020;10(November):1–17.
16. Huang Y, Yang C, Xu X-F, Xu W, Liu S-W. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. 2020; Available from: <https://doi.org/10.1038/s41401-020-0485-4>
17. Mengist HM, Kombe Kombe AJ, Mekonnen D, Abebaw A, Getachew M, Jin T. Mutations of SARS-CoV-2 spike protein: Implications on immune evasion and vaccine-induced immunity. *Semin Immunol* [Internet]. 2021;55(January):101533. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/34836774>
18. Malik YA. Properties of coronavirus and SARS-CoV-2. *Malays J Pathol.* 2020;42(1):3–11.
19. Xia X. Domains and functions of spike protein in sars-cov-2 in the context of vaccine design. *Viruses.* 2021 Jan 1;13(1).
20. Parasher A. COVID-19: Current understanding of its Pathophysiology, Clinical presentation and Treatment. *Postgrad Med J.* 2021;97(1147):312–20.
21. Triggler CR, Bansal D, Ding H, Islam MM, Farag EABA, Hadi HA, et al. A Comprehensive Review of Viral Characteristics, Transmission, Pathophysiology, Immune Response, and Management of SARS-CoV-2 and COVID-19 as a Basis for Controlling the Pandemic. *Front Immunol.* 2021;12(February):1–23.
22. Fernandes APM, Maciel LMZ, Foss MC, Donadi EA. Como entender a associação entre o sistema HLA e as doenças auto-imunes endócrinas. *Arq Bras Endocrinol Metabol.* 2003;47(5):601–11.
23. Luring AS, Hodcroft EB. Genetic Variants of SARS-CoV-2 - What Do They Mean? *JAMA - J Am Med Assoc.* 2021;325(6):529–31.
24. Shiehzadegan S, Alaghemand N, Fox M, Venketaraman V. Analysis of the Delta Variant B.1.617.2 COVID-19. *Clin Pract* [Internet]. 2021 Oct 21;11(4):778–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/34698149>
25. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell.* 2020;181(7):1489-1501.e15.
26. NetCTLpan [Internet]. Available from: <https://services.healthtech.dtu.dk/service.php?NetCTLpan-1.1>
27. Baral P, Bhattarai N, Hossen ML, Stebliankin V, Gerstman BS, Narasimhan G, et al. Mutation-induced changes in the receptor-binding interface of the SARS-CoV-2 Delta variant B.1.617.2 and implications for immune evasion. *Biochem Biophys Res Commun* [Internet]. 2021;574:14–9. Available from: <https://doi.org/10.1016/j.bbrc.2021.08.036>
28. Kojima N, Shrestha NK, Klausner JD. A Systematic Review of the Protective Effect

of Prior SARS-CoV-2 Infection on Repeat Infection. *Eval Heal Prof.*
2021;44(4):327–32.

APÊNDICE

A. POTENCIAS EPÍTOPOS ENCONTRADOS PELO NETCTLPAN PARA HLA-A, HLA-B E HLA-C DE PROTEÍNA SPIKE REFERENTE À SEQUÊNCIA DE REFERÊNCIA

Sequência referência - Spike						
Peptídeos que se ligam ao HLA A com rank ≤ 0.15		Peptídeos que se ligam ao HLA B com rank ≤ 0.15		Peptídeos que se ligam ao HLA C com rank ≤ 0.15		
	Posição	Quantidade	Posição	Quantidade	Posição	Quantidade
1	360	10	868	7	868	15
2	161	7	895	6	885	9
3	257	7	83	5	1	7
4	268	6	159	4	159	7
5	301	6	235	4	23	4
6	453	6	711	4	101	4
7	1098	6	713	4	365	4
8	88	4	814	4	368	4
9	191	4	1211	4	504	3
10	457	4	1015	3	895	3
11	603	4	23	2	161	2
12	1236	4	55	2	257	2
13	717	3	154	2	717	2
14	732	3	1180	2	195	1
15	864	3	7	1	83	1
16	1207	3	161	1	192	1
17	35	2	167	1	201	1
18	369	2	240	1	268	1
19	416	2	344	1	356	1
20	724	2	368	1	508	1
21	776	2	463	1	686	1
22	868	2	556	1		
23	975	2	589	1		
24	1059	2	624	1		
25	143	1	698	1		
26	201	1	877	1		
27	356	1	1205	1		
28	365	1				
29	488	1				
30	634	1				
31	690	1				
32	733	1				
33	1064	1				

B. POTENCIAS EPÍTOPOS ENCONTRADOS PELO NETCTLPAN PARA HLA-A, HLA-B E HLA-C NA SEQUÊNCIA DE PROTEÍNA SPIKE DA VARIANTE DELTA

EPI_ISL_6661895 - Delta 1 - Spike						
Peptídeos que se ligam ao HLA A com rank ≤ 0.15			Peptídeos que se ligam ao HLA B com rank ≤ 0.15		Peptídeos que se ligam ao HLA C com rank ≤ 0.15	
	Posição	Quantidade	Posição	Quantidade	Posição	Quantidade
1	358	10	866	7	866	15
2	159	7	893	6	883	9
3	255	7	83	5	1	7
4	299	6	149	4	157	7
5	451	6	157	4	23	4
6	1096	6	233	4	101	4
7	266	5	709	4	363	4
8	88	4	711	4	366	4
9	149	4	812	4	502	3
10	189	4	1209	4	893	3
11	455	4	1013	3	159	2
12	601	4	23	2	238	2
13	1234	4	55	2	255	2
14	715	3	152	2	715	2
15	730	3	1178	2	83	1
16	862	3	7	1	199	1
17	35	2	159	1	266	1
18	143	2	165	1	506	1
19	367	2	238	1	684	1
20	414	2	342	1		
21	722	2	366	1		
22	774	2	461	1		
23	866	2	587	1		
24	973	2	622	1		
25	1057	2	696	1		
26	199	1	875	1		
27	354	1	1203	1		
28	363	1				
29	486	1				
30	632	1				
31	688	1				
32	731	1				
33	1062	1				
34	1205	1				