

**CENTRO UNIVERSITÁRIO DE BRASÍLIA  
FACULDADE DE CIÊNCIAS DA EDUCAÇÃO E SAÚDE  
GRADUAÇÃO EM BIOMEDICINA**

**RORDANA GOMES FERNANDES PEREIRA**

**ANÁLISE MOLECULAR, BIOLÓGICA E FILOGENÉTICA DO NOVO  
CORONAVÍRUS**

Trabalho de conclusão de curso apresentado em forma de artigo como requisito, do Curso de Biomedicina do Centro Universitário de Brasília (CEUB), sob orientação do Prof. Paulo Roberto Martins Queiroz.

Brasília

2021

## **Análise molecular, biológica e filogenética do novo Coronavírus**

Rordana Gomes Fernandes Pereira<sup>1</sup>

Paulo Roberto Martins Queiroz<sup>2</sup>

### **Resumo**

Os Coronavírus são uma família de vírus que possuem RNA de fita simples e são capazes de contaminar homens e animais. São conhecidos por sua alta frequência de recombinação e altas taxas de mutação, permitindo que se adaptem a novos hospedeiros e nichos ecológicos. Os níveis virais elevados fornecem um ambiente adequado para a replicação viral, levando a persistência viral, por sua vez, pode ocasionar novas estirpes virais. Foi realizada uma revisão bibliográfica narrativa sobre a análise molecular e biológica. Na análise filogenética, executou-se uma pesquisa em base de dados genéticos. A infecção por este vírus encontra populações suscetíveis em todo o mundo, e à medida que a taxa de replicação viral aumenta nessas populações, o número de novos pontos de mutação e as frequências de mutações em pontos de acesso individuais aumentaram. Assim, pode tornar seu genoma mais bem adaptado e cada vez mais perigoso para a saúde humana.

**Palavras-chave:** Coronavírus nos humanos; SARS-CoV-2; Doença do Coronavírus 2019; Coronavírus; Genômica viral.

### **Molecular, biological and phylogenetic analysis of the new Coronavirus**

#### **Abstract**

Coronaviruses are a family of viruses that have single-stranded RNA and are capable of infecting humans and animals. They are known for their high recombination frequency and high mutation rates, allowing them to adapt to new hosts and ecological niches. High viral levels provide a suitable environment for viral replication, leading to viral persistence, which in turn can give rise to new viral strains. A narrative literature review on molecular and biological analysis was carried out. In the phylogenetic analysis, a search was carried out in a genetic database. Infection with this virus finds susceptible populations around the world, and as the rate of viral replication in these populations increases, the number of new mutation points and the frequencies of mutations at individual hotspots have increased. Thus, it can make its genome better adapted and increasingly dangerous to human health.

**Keywords:** Coronavirus in humans; SARS-CoV-2; Coronavirus disease 2019; Coronavirus; Viral genomics.

---

<sup>1</sup> Graduando de Biomedicina do CEUB

<sup>2</sup> Professor do curso de Biomedicina no CEUB

## 1. INTRODUÇÃO

Os Coronavírus (CoVs) pertencem a família *Coronaviridae* composta por vírus que possuem RNA de fita simples e são capazes de contaminar homens e animais, ocasionando doenças respiratórias, hepáticas, gastrointestinais e neurológicas. A família é dividida em quatro gêneros: alfa-coronavírus, beta-coronavírus, gama-coronavírus e delta-coronavírus (WU, A. et al., 2020). Atualmente, foram reconhecidos seis Coronavírus humanos: alfa-CoVs HCoV-NL63 e HCoV-229E e os beta-CoVs HCoV-OC43, HCoV-HKU1, síndrome respiratória aguda grave-CoV (SARS-CoV) (DROSTEN et al., 2003) e síndrome respiratória do Oriente Médio-CoV (ZAKI et al., 2012).

A coevolução dos CoVs e seus hospedeiros, incluindo os seres humanos, foi impulsionada pela mutação e adaptação. Antes de 2003, ficou conhecido por causar doenças leves, como o resfriado comum. Após isso, surgiram os surtos da síndrome respiratória aguda grave (SARS) e da síndrome respiratória do Oriente Médio (MERS) que revelaram o quão devastador e potencialmente fatal uma infecção por HCoV pode ser (GAO et al., 2016).

O surgimento do SARS-CoV-2 na China no final de 2019 colocou os CoVs, novamente, no centro das atenções devido sua alta capacidade de transmissibilidade. A identificação dos hospedeiros animais têm interferências diretas em relação a prevenção de doenças humanas. A partir disso, a maioria dos HCoVs se originou de morcegos, onde eles não são patogênicos (YE et al., 2020).

O novo Coronavírus é suspeito de ter origem na vida selvagem, com origens finais prováveis em uma espécie de morcego e possivelmente com origens mais proximais de um hospedeiro animal intermediário. A partir da propagação global do vírus entre os humanos, surge a preocupação que este novo vírus tenha o potencial de se espalhar de humanos infectados para a vida selvagem nativa que poderia posteriormente servir como novos hospedeiros reservatórios para o vírus. Dessa forma, o vírus pode se estabelecer em áreas fora de sua região de origem e ficar disponível para surtos futuros (JI et al., 2020).

Devido à alta taxa de prevalência e extensa distribuição dessa família de vírus e, além disso, a grande diversidade genética e recombinação frequente de seus genomas e, também, ao aumento das atividades de interação humano-animal promovem, principalmente, o surgimento de novos Coronavírus em humanos (WU, A. et al., 2020). Ademais, estruturalmente, eles possuem genomas não segmentados que compartilham uma organização semelhante. E,

também, o período de incubação e a duração do curso da doença pelo HCoV são muito parecidos (WONG; LUI; JIN, 2016).

O vírus se espalhou globalmente, e a COVID-19 foi declarado pela Organização Mundial da Saúde (OMS) uma pandemia em março de 2020. Por ser uma doença infecciosa respiratória aguda emergente ela se espalha, principalmente, através do trato respiratório por secreções respiratórias, contato direto e gotículas. Ademais, foi descrito que o SARS-CoV-2 foi isolado de esfregaços fecais e sangue, demonstrando a possibilidade de múltiplas rotas de transmissão. Informações atuais sugerem um período de incubação de 1–14 dias, na maioria dos casos 3–7 dias. É um vírus responsável por causar problemas graves, especialmente em idosos e pessoas com doenças crônicas subjacentes (LUDWIG; ZARBOCK, 2020).

Os pacientes com COVID-19, na maioria, apresentam uma doença leve que pode ser tratada na comunidade. Porém, eles requerem monitoramento cuidadoso e encaminhamento precoce ao hospital em caso de qualquer sinal de agravamento clínico. Sendo que, atualmente, o protocolo médico usado nos pacientes é predominantemente de suporte, com foco no manejo adequado da disfunção respiratória (THEVARAJAN; BUISING; COWIE, 2020).

Ademais, o reconhecimento da infecção assintomática tem sido uma área de intenso interesse na compreensão da epidemiologia da nova doença. A proporção de infecção sintomática e assintomática é, atualmente, incerta (SUTTON et al., 2020). A forma sintomática da infecção geralmente se apresenta como uma síndrome respiratória, mais frequentemente com febre e tosse (WANG et al., 2020).

Os sintomas clínicos mais comuns são febre e tosse, falta de ar e outras dificuldades respiratórias, além de outros sintomas inespecíficos, incluindo dor de cabeça, dispneia, fadiga e dor muscular. Além disso, alguns pacientes também relatam sintomas digestivos, como diarreia e vômitos. Embora os pacientes inicialmente apresentem febre com ou sem sintomas respiratórios, vários graus de anormalidades pulmonares se desenvolvem posteriormente em todos os pacientes, e podem ser observados na tomografia computadorizada (TC) de tórax (OUASSOU et al., 2020).

Contudo, informações atualizadas sobre genômica viral e transcriptômica são cruciais para compreender as origens e a dispersão global do vírus. Com isso, possibilitando estratégias para intervenções terapêuticas, descoberta de medicamentos e desenvolvimento de vacinas (UDDIN et al., 2020).

A partir disso, o primeiro genoma SARS-CoV-2 foi publicado apenas algumas semanas após o início do surto, e exibiu similaridade genômica e filogenética com o SARS-CoV,

particularmente no gene S e domínio de ligação ao receptor (RBD), indicando a capacidade de transmissão direta de pessoa para pessoa (ANDERSEN et al., 2020).

Com base em suas relações filogenéticas e estruturas genômicas, o COVID-19 pertence ao gênero *Betacoronavirus*. Os Betacoronavírus humanos (SARS-CoV-2, SARS-CoV e MERS-CoV) têm muitas semelhanças, mas também têm diferenças em sua estrutura genômica e fenotípica que podem influenciar sua patogênese (MOUSAVIZADEH; GHASEMI, 2020).

As análises filogenéticas dos genomas SARS-CoV-2 identificaram os morcegos como o reservatório primário de Coronavírus semelhantes ao SARS, exibindo alta similaridade de sequência (96,2%) entre os genomas BatCoV e SARS-CoV-2. A análise de sequência da proteína spike sugere ainda que novas mutações em seu RBD determinam não apenas a gama de hospedeiros, mas também o tropismo celular do vírus (CERAOLO; GIORGI, 2020).

Além disso, uma compreensão dos reservatórios biológicos que transportam Coronavírus e as modalidades de contato com a população humana através do comércio, viagens ou recreação será importante para compreender os riscos futuros de novas infecções. Além do que, as populações podem ser infectadas ou mesmo reinfectadas por meio de várias rotas de viagem (KOYAMA; PLATT; PARIDA, 2020).

Em muitas partes do mundo houve o sequenciamento do novo vírus, e esses dados provaram ser úteis no rastreamento da disseminação global do vírus (HADFIELD et al., 2018). Por exemplo, uma análise inicial de 103 genomas SARS-CoV-2 identificou dois subtipos principais (que foram designados L e S) que são bem definidos por dois polimorfismos de nucleotídeo único (SNPs) diferentes (TANG et al., 2020).

Os vírus de RNA tendem a abrigar polimerases de RNA dependentes de RNA, propensas a erros, que tornam a ocorrência de mutações e eventos de recombinação bastante frequentes. Sendo assim, pode desempenhar um papel na evolução do SARS-CoV-2 (DUFFLY, 2018).

Em Wuhan, China, o tipo L foi encontrado em 70% dos casos e foi observada como a forma mais agressiva e contagiosa em comparação com o tipo S original. O vírus sofreu mutação adicional e se expandiu em várias cepas e grupos. A diversidade geográfica de diferentes cepas pode ajudar a correlacionar a gravidade relacionada a COVID-19, a taxa de mortalidade e as opções de tratamento (FORSTER et al., 2020).

Assim, este trabalho tem como objetivo analisar o aspecto molecular, biológico e filogenético do SARS-CoV-2. Para isso, foi feito um levantamento dos estudos relacionados com o tema e utilização de um software (em relação a análise filogenética) para elaboração do

mesmo, com a finalidade de verificar e compreender o comportamento viral nos países selecionados para o estudo.

## 2. MÉTODO

O trabalho foi dividido em análise biológica, molecular e filogenética do novo Coronavírus. Para isso, foi realizada uma revisão bibliográfica narrativa em relação a análise biológica e molecular. Em complemento, executou-se a análise filogenética em base de dados genéticos.

Para a realização e discussão do trabalho foi feito um levantamento bibliográfico a respeito de publicações relacionadas ao Coronavírus, principalmente, nos últimos 2 anos. As palavras-chave utilizadas na elaboração do trabalho foram: Coronavírus nos humanos; SARS-CoV-2; Doença do Coronavírus 2019; Coronavírus; Genômica viral utilizando a base bibliográfica do PubMed (Public Medline) com a seleção de artigos em língua inglesa relacionados com o tema. Além disso, foi utilizada apenas uma referência da SciELO em língua portuguesa.

Na análise filogenética, foram utilizados 105 genomas selecionados de forma aleatória pela plataforma Genbank. Essa plataforma é um banco de dados de nucleotídeos do NCBI (*National Center for Biotechnology Information* - Centro Nacional de Informações sobre Biotecnologia), dessa forma, armazenando informações sobre sequências nucleotídicas.

A partir da seleção das sequências elas foram inseridas no programa Geneious, um software que estrutura as sequências genéticas e faz o alinhamento e a análise da identidade das sequências de DNA. Com isso, árvores filogenéticas foram criadas de acordo com as sequências de DNA dos genomas virais que foram selecionadas e organizadas segundo cada país com as sequências WIV (são sequências de Coronavírus identificadas com provável origem em morcegos hospedeiros) e, por fim, uma árvore contendo 105 genomas, sendo: 19 sequências obtidas de hospedeiros humanos no Brasil, 21 sequências obtidas de amostras humanas provenientes da China, 20 sequências originárias da Índia, 20 sequências provenientes da Itália, 20 sequências obtidas de amostras dos Estados Unidos e 5 sequências com provável origem de morcegos (WIV).

As sequências isoladas de humanos são relacionadas com as sequências WIV devido elas serem as primeiras sequenciadas após a transmissão do Coronavírus para humanos com provável origem de morcego. A WIV04, por exemplo, foi a primeira isolada em humanos.

Essas sequências foram selecionadas para o trabalho com a finalidade de analisar a filogenética do SARS CoV-2 em relação às mutações a partir das sequências WIV. Sendo que, as sequências WIV (*Wuhan Institute of Virology* - Instituto Wuhan de Virologia) representam as sequências publicadas no Genbank devido a um surto de pneumonia associado a um novo Coronavírus de provável origem do morcego. Por isso, foi escolhido e inserido na pesquisa para contribuir na análise de filogenia e verificar, dessa forma, o comportamento viral de mutação em cada nacionalidade.

### **3. RESULTADOS E DISCUSSÃO**

#### **3.1 Análise Molecular do SARS CoV-2**

A epidemia de síndrome respiratória aguda grave (SARS) na China, que matou 774 pessoas durante o surto de 2002/2003 foi a primeira doença grave conhecida causada por um Coronavírus (CoV) em humanos. A segunda doença fatal foi a síndrome respiratória do Oriente Médio (MERS) que teve início em 2012 na Arábia Saudita e resultou em 858 mortes. Em dezembro de 2019, o SARS-CoV-2 (COVID-19), um novo vírus, com origem da China, começou a gerar repercussão em todo o mundo devido à velocidade sem precedentes de sua transmissão. Foram 5,2 milhões de pessoas infectadas e 338.480 óbitos de dezembro de 2019 a maio de 2020 (LATIF; MUKARATIRWA, 2020).

Acredita-se que esses Coronavírus humanos sejam de origem animal e atingiram os humanos através do salto de espécies. Eles são bem conhecidos por sua alta frequência de recombinação e altas taxas de mutação, permitindo que se adaptem a novos hospedeiros e nichos ecológicos (CUI; LI; SHI, 2019).

O primeiro reconhecimento do vírus em humanos ocorreu em 1960, sendo que sete tipos foram identificados (SEAH; AGRAWAL, 2020). Dessa forma, eles são classificados com base em sua sequência de proteínas em: Alphacoronavirus, Betacoronavirus, Gammacoronavirus e Deltacoronavirus. O Betacoronavírus e o alfa são os dois gêneros que infectam mamíferos, enquanto o gamma e delta infectam aves (WU, D. et al., 2020).

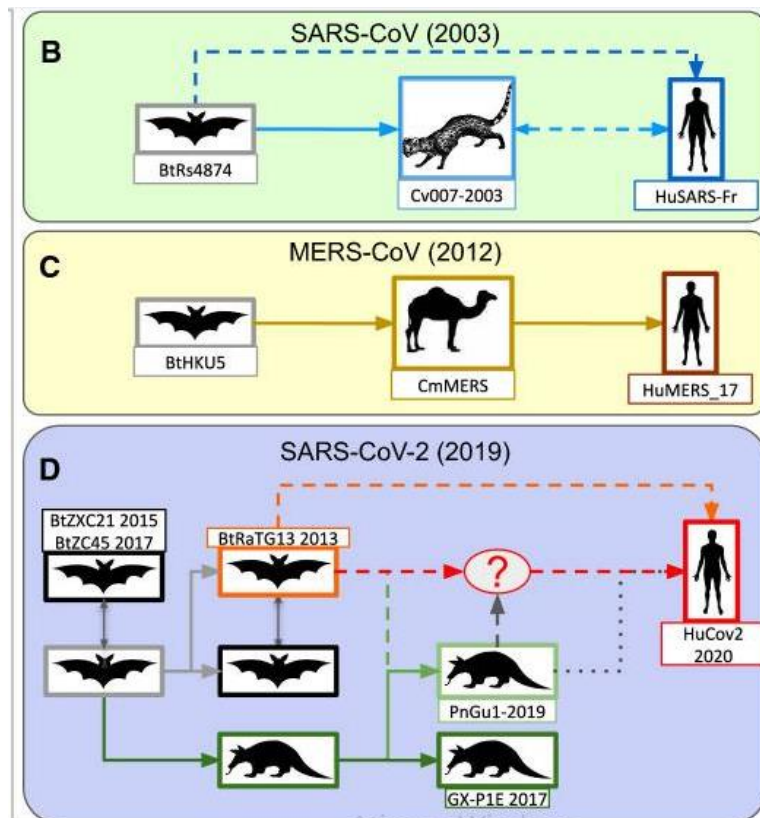
Em geral, os vírus de RNA codificam para proteínas que são fundamentais para controlar as células hospedeiras. Por exemplo, eles codificam os tipos de proteínas necessários para montar as novas partículas virais. Além do que, os vírus de RNA são bem conhecidos por

sua alta taxa de mutação, o que os ajuda a se adaptarem rapidamente às células hospedeiras (MANFREDONIA et al., 2020).

Os CoVs são uma família de vírus que infectam mais de 500 espécies de quirópteros (uma ordem de mamíferos composta por mais de 1200 espécies de morcegos) que representam um importante reservatório para a evolução do CoV, permitindo a recombinação de genomas virais em animais co-infectados por diferentes estirpes (HU et al., 2017; LUK et al., 2019; MENACHERY et al., 2015).

É geralmente aceito que a transmissão zoonótica de CoVs para humanos ocorre por meio de uma espécie hospedeira intermediária, na qual os vírus melhor adaptados aos receptores humanos podem ser selecionados, facilitando assim a passagem da barreira de espécies (CUI; LI; SHI, 2019). Vetores de transmissão zoonótica podem ser identificados examinando as relações filogenéticas entre novos vírus e vírus isolados de espécies animais que vivem nas regiões de emergência do CoV (SALLARD et al., 2021). Na figura 1, é possível verificar hipóteses de transmissão do reservatório animal (morcegos) para o homem, com base na filogenia molecular dos genomas virais.

**Figura 1:** Esquemas do possível reservatório animal para o homem com base na filogenia molecular.



Fonte: Sallard et al. (2021).



Em relação a figura 1, a pandemia de SARS-CoV de 2003, o civeta (gatos selvagens) foi proposto como hospedeiro intermediário. A transmissão direta de morcego para humano também está sendo considerada. No caso de MERS-CoV de 2012, com o camelo como hospedeiro intermediário. Vários eventos de transmissão direta foram documentados. E na pandemia de COVID-19 vários cenários são propostos sobre o último hospedeiro antes da transmissão aos humanos (SALLARD et al., 2021).

Embora nenhuma epidemia relacionada à transmissão direta de morcego para humano tenha sido identificada até o momento, estudos experimentais mostraram que mais de 60 CoVs de quirópteros são capazes de infectar células humanas em cultura (LUIS et al., 2013; MENACHERY et al., 2015). A identificação, em 2017, de isolados virais muito semelhantes ao SARS-CoV em morcegos levanta a questão de uma possível transmissão direta de quirópteros para humanos, que poderia resultar de mutações no domínio de ligação ao receptor (RBD) da proteína spike viral e permitiu sua entrada na célula hospedeira (HU et al., 2017).

Além disso, a origem do SARS-CoV-2 é uma questão de debate. Estudos de bioinformática revelaram que ele tem 96,2% de identidade com um genoma CoV (RaTG13) reconstruído a partir de fezes e amostras anais de morcegos *Rhinolophus affinis*. Curiosamente, essas amostras foram coletadas em 2013, mas a sequência do genoma completo só foi publicada no início de fevereiro de 2020 (ZHOU et al., 2020).

Infelizmente, a localização precisa da coleta de amostra não está documentada no artigo original nem nos bancos de dados de sequência. No entanto, quando a análise atual foi conduzida (abril de 2020), foi encontrada uma correspondência exata entre RaTG13 e um fragmento de 370 nucleotídeos publicado em 2016 (KP876546), codificando um domínio de polimerase BtCoV/4991, que foi sequenciado de isolados coletados de um mina eixo na província de Yunnan após a morte de 3 mineiros de pneumonia atípica (GE et al., 2016; WU et al., 2014; RISSANEN et al., 2017).

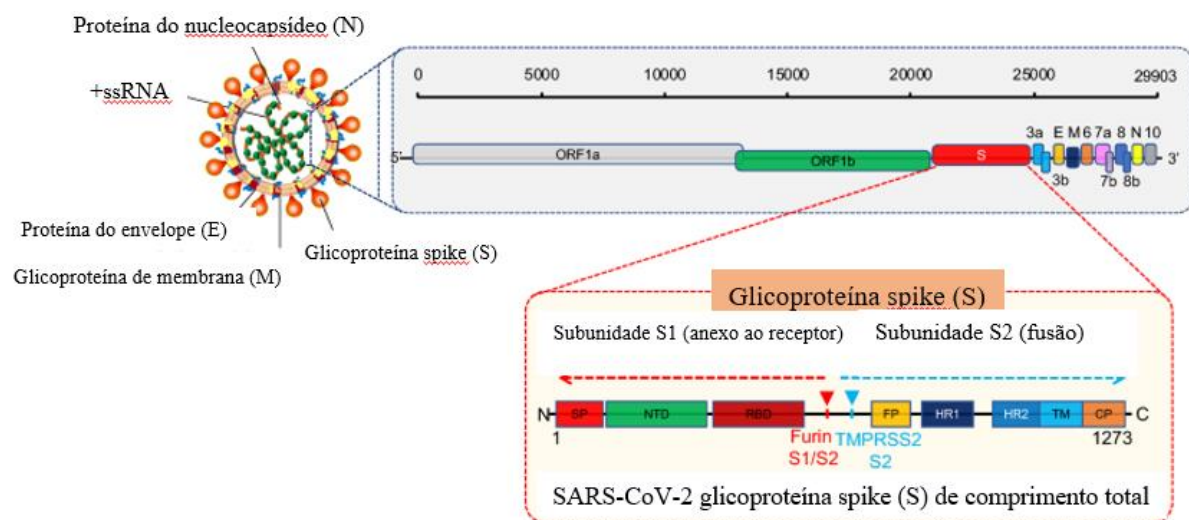
O genoma dos CoVs representa o segundo maior genoma de todos os vírus de RNA. É um RNA de fita simples de sentido positivo, com tamanho do genoma variando de 26 kb a 32 kb de comprimento e tem um cap e poli-A em sua cauda 5' e 3', respectivamente (CHAN et al., 2020, WU et al., 2020). Em geral, o número de ORFs no genoma dos CoVs varia de 6 a 15 ORFs (SONG et al., 2019).

As ORFs 1a e 1ab representam o maior gene do genoma do Coronavírus e cobrem quase dois terços de todo o seu genoma, como pode ser observado na figura 2. Esses dois genes codificam diferentes NSPs como o complexo de replicação-transcrição (RTC), que é

responsável por sintetizar e transcrever o RNA subgenômico (sgRNA) (GORBALENYA et al., 2000; HUSSAIN et al., 2005; SNIJDER et al., 2006).

Assim, o genoma do novo vírus consiste em 15 ORFs codificados em 29 proteínas (SRINIVASAN et al., 2020); no terminal 5' do genoma, temos ORF1ab e ORF1a que codificam o polipeptídeo 1ab e 1a, respectivamente. O terminal 3' do genoma representa quatro proteínas estruturais (pico, envelope, matriz e nucleocapsídeo) e nove acessórias (3a, 3b, 6, 7a, 7b, 8b, 9a, 9b e orf10) (AL-QAANEH et al., 2021).

**Figura 2:** Estrutura viral do SARS CoV-2.



Fonte: Adaptado de Kumar e Khodor (2020).

Sendo que, as duas estruturas de leitura aberta sobrepostas (ORFs) ORF1a e 1ab representam os dois terços da região a montante do genoma. A característica genética geral que distingue o Alphacoronavirus de outros Coronavírus é que eles têm um tipo único de proteína não estrutural NSP1 que difere em sequência e tamanho de outro Betacoronavirus NSP1 (~9 kDa para Alphacoronavirus e ~20 kDa para Betacoronavirus) (AL-QAANEH et al., 2021).

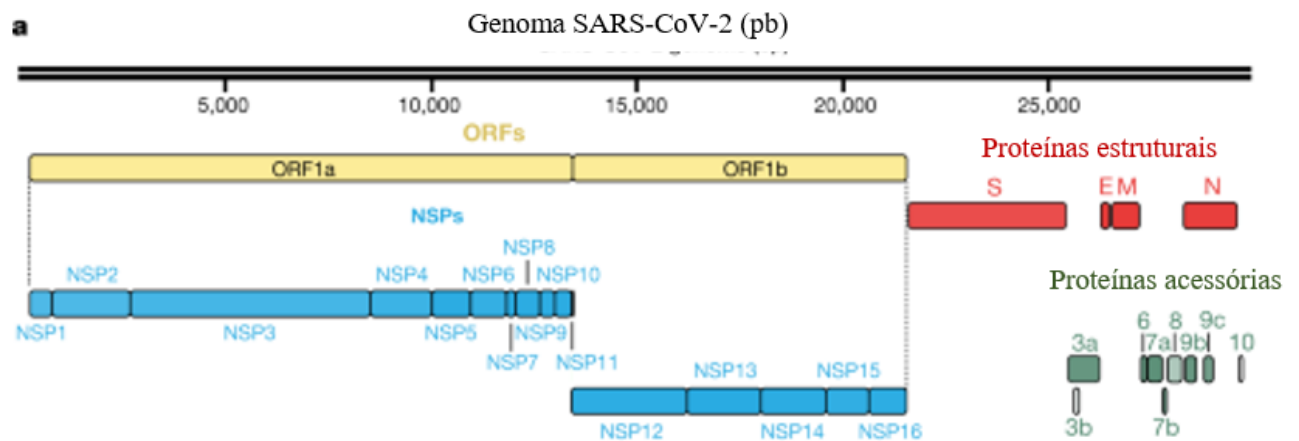
Em contraste, Gamma e Deltacoronavírus carecem da porção NSP1 (JAIMES et al., 2020). A proteína NSP1 desempenha um papel crucial na inibição da expressão do gene do hospedeiro e sua localização é essencial para a virulência do vírus (SHEN et al., 2019). Conforme a figura 3, é possível observar a organização genômica do SARS CoV-2.

A partir disso, esse vírus contém quatro proteínas estruturais (S, E, M e N) e dezesseis proteínas não estruturais (nsp1 a 16) que estão dispostas de acordo com a figura 3. A Nsp1 medeia o processamento e a replicação do RNA. Enquanto a Nsp2 modula a via de sinalização

de sobrevivência da célula hospedeira. Acredita-se que o Nsp3 separa a proteína traduzida. E a Nsp4 contém o domínio transmembranar 2 (TM2) e modifica as membranas ER. Já a Nsp5 participa do processo de poliproteína durante a replicação. A Nsp6 é um domínio transmembranar presuntivo. A presença de nsp7 e nsp8 aumentou significativamente a combinação de nsp12 e RNA template-primer (NAQVI et al., 2020).

Além disso, a Nsp9 funciona como uma proteína de ligação a ssRNA. A Nsp10 é crítica para a metilação cap de mRNAs virais. Enquanto a Nsp12 contém a RNA polimerase dependente de RNA (RdRp), que é uma composição crítica da replicação/transcrição do coronavírus. O Nsp13 liga-se ao ATP e o domínio de ligação do zinco no Nsp13 participa do processo de replicação e transcrição. Já a Nsp14 é um domínio de exoribonuclease de revisão. Enquanto a Nsp15 tem atividade de endoribonuclease dependente de Mn (2+). E, por fim, a Nsp16 é uma metiltransferase 2'-O-ribose (NAQVI et al., 2020).

**Figura 3:** Organização genômica do SARS CoV-2.



Fonte: Adaptado de Gordon et al. (2020).

Um estudo mostra que existem alguns efeitos mediados por NSP no *splicing*, tradução e tráfego de proteínas para inibir as defesas do hospedeiro. Após a infecção por SARS-CoV-2, a NSP16 se liga a domínios de reconhecimento de mRNA dos snRNAs U1 e U2 para suprimir o *splicing* de mRNA. O NSP1 se liga ao RNA ribossômico 18S no canal de entrada do mRNA do ribossomo para interferir na tradução do mRNA. NSP8 e NSP9 se ligam ao RNA 7SL que se localiza na Partícula de Reconhecimento de Sinal para interromper o tráfego de proteínas para a membrana celular (BANERJEE et al., 2020).

A sequência reguladora da transcrição que medeia o processo de transcrição está localizada entre ORFs no sgRNA e serve como um modelo para a síntese de mRNA (SAWICKI

et al., 2007). A mutação *Frameshift* (mudança na matriz de leitura) entre ORF1a e ORF1b resulta na síntese de dois polipeptídeos (pp1a e pp1ab) que serão então processados em 16 NSPs com o auxílio de proteases semelhantes à quimotripsina (3CLpro) ou proteases semelhantes à papaína com a protease principal (Mpro) (MASTERS, 2006).

O terço restante do genoma do CoV é responsável por codificar pelo menos quatro proteínas estruturais como proteína spike (S), proteína do nucleocapsídeo (N), proteína do envelope (E) e proteínas de membrana (M), além de algumas proteínas acessórias, como como proteínas 3a/b, 4a/b e Hemaglutinina-Esterase (HE) (HUSSAIN et al., 2005).

O alinhamento de sequência do genoma de todos os CoVs ilustra uma identidade de 43% para regiões codificantes de proteínas estruturais e 58% para regiões codificantes de proteínas não estruturais. Em comparação, a identidade em todo o genoma entre todos os CoVs é de cerca de 54%. Esses resultados sugerem que as proteínas estruturais têm mais diversidade do que as outras proteínas não estruturais (CHEN et al., 2020).

A proteína não estrutural 14 (NSP-14) é uma enzima exoribonuclease 3' a 5' distinta para todos os CoVs. Sua função está relacionada à manutenção de todo o genoma do RNA dos CoVs e à revisão do complexo replicação-transcrição (ECKERLE et al., 2010, SMITH et al., 2013). Verificou-se que as sequências de aminoácidos das proteínas N e E no SARS-CoV-2 têm uma identidade de sequência de ~92% (AZEEZ et al., 2020).

Uma vez que o genoma do novo Coronavírus tem a capa 5' metilada (onde as proteínas não estruturais são colocadas) e o lado 3' poliadenilado (onde residem as proteínas estruturais), isso permite a tradução direta após a infecção sem a necessidade de um estágio de transcrição intermediário. Além disso, o genoma SARS-CoV-2 inclui vários ORFs que podem ser transcritos por várias sequências reguladoras da transcrição (TRS). O ciclo de vida viral começa quando o vírion entra no trato respiratório humano e depende da interação do vírus com a membrana da célula epitelial. A infecção inicia na superfície celular quando a subunidade S1 reconhece a enzima conversora de angiotensina 2 (ECA2) como um receptor-alvo ou reconhece as exopeptidases como um receptor-chave para a entrada na célula hospedeira. O mecanismo de entrada depende de proteases celulares, como catepsinas e protease transmembrana serina 2 (TMPRSS2), e tripsina protease humana das vias aéreas (HAT) (HOFFMANN et al., 2020).

A subunidade S1 consiste em dois subdomínios que podem operar como proteínas de domínios de ligação ao receptor (RBD); Domínio N-terminal (NTD) como uma ligação de açúcar mediada e um domínio C-terminal (CTD) como reconhecimento de receptores (XIA et al., 2020). O RBD mutado da subunidade S1 é necessário para espécies cruzadas de transmissão

de SARS-CoV-2. Uma vez que a ligação ocorre, mudanças conformacionais na proteína S, como redução do pH e/ou proteólise da proteína S, promovem a separação de S1 de S2 (AL-QAANEH et al., 2021).

Conseqüentemente, a fusão da membrana é colocada pela região S2 (HAN et al., 2020). A subunidade S2 tem dois domínios de repetição de heptal (HR), (HR1) e (HR2), que interagem juntos para formar um núcleo de fusão de feixe de seis hélices (6-HB) para aproximar a interação viral e celular (XIA et al., 2020).

Os genes da replicase/transcriptase codificados por ORF1a/1ab traduzem duas poliproteínas (pp1a e pp1ab) que são clivadas em 16 proteínas não estruturais individuais (nsp1 a 16) pela ação da enzima protease de cisteína semelhante à quimotripsina (3CLpro, nsp3) que controlam a replicação viral e é crucial para o ciclo de vida do SARS-CoV-2 e a enzima protease semelhante à papina (PLpro, protease nsp5) que tem um papel crucial na supressão do sistema imunológico por desubiquitinação das proteínas da célula hospedeira (MORSE et al., 2020; QAMAR et al., 2020).

A RNA polimerase dependente de RNA (RdRp) que é codificada por NSP12 desempenha um papel central na síntese de uma nova molécula do genoma do vírus (replicação). Também é responsável pela síntese de modelos subgenômicos (constituição de RNAs da estrutura da partícula viral do SARS-CoV-2) para produção de RNA mensageiro (transcrição) (GAO et al., 2020).

As proteínas transcritas são inseridas no retículo endoplasmático e o aparelho de Golgi traduz o RNA compactado dentro do capsídeo formado. A proteína N forma o capsídeo, enquanto o envelope, a membrana e a spike são formados pelas proteínas E, M e S, respectivamente. Finalmente, o vírus das vesículas compactadas é liberado da membrana celular pelo processo de exocitose (SHEREEN et al., 2020).

### **3.2. Análise Biológica do SARS-CoV-2**

Os Coronavírus, recentemente, retornaram com um novo, o SARS-CoV-2, relacionado a uma doença respiratória potencialmente grave. A condição tem gerado preocupações consideráveis devido à sua rápida e ampla distribuição (NASAB; SAGHAZADEH; REZAEI, 2020).

O transbordamento zoonótico, também conhecido como transmissão de animal para humano, é referido como quando um patógeno é transmitido de um animal não humano para um humano. Dessa forma, é o resultado de processos contínuos de modificações de um

patógeno animal para se tornar um zoonótico. Com isso, esses processos contribuem para que um patógeno animal possa lidar com os obstáculos que naturalmente enfrentará ao se tornar um patógeno zoonótico. Sendo que, os obstáculos se relacionam, principalmente, com características específicas do patógeno (prevalência, liberação e sobrevivência), a exposição e as interações reservatório-hospedeiro (PLOWRIGHT et al., 2017).

Um estudo revela que o SARS CoV-2 pode ser agrupado com o Coronavírus tipo SARS de morcego isolado de 2015. A pesquisa defende a hipótese de que o novo vírus provavelmente se originou de morcegos e atingiu os humanos. Assim, a análise da estrutura da proteína mostra que a proteína do nucleocapsídeo (N) e nucleoproteína semelhante a glicoproteína spike (S) no SARS CoV-2 compartilham semelhanças significativas com as do Coronavírus SARS semelhante ao morcego. No entanto, existem algumas diferenças que, dada a importância dessas duas proteínas na eficiência da montagem viral e na fusão da membrana celular, podem ser responsáveis pelo aumento da patogenicidade do novo Coronavírus em comparação com a da SARS semelhante ao morcego (BENVENUTO et al., 2020).

Por outro lado, pesquisas recentes sugerem um possível hospedeiro intermediário. O Pangolin-CoV, encontrado nos pangolins malaios, compartilha 91,02% de similaridade com o SARS Cov-2 no nível de genoma. Entretanto, nem Bat-CoV nem Pangolin-CoV possuem um local de clivagem de furina na margem S1/S2 da glicoproteína S, diferenciando-os do SARS CoV-2 (ZHANG; WU; ZHANG, 2020).

Além do que, os morcegos podem desempenhar um papel como hospedeiro reservatório para uma variedade de vírus devido ao seu sistema imunológico exclusivo. Os interferons (INs) são os agentes antivirais imunomoduladores que podem possibilitar a fagocitose por macrófagos e conter a propagação e inibir a replicação do vírus. Logo, a inibição da produção de IFN enfraqueceria a primeira linha de defesa e, assim, ajudaria na sobrevivência do vírus (LI, D. et al., 2020).

Além do mais, não apenas os morcegos podem regular a resposta do IFN tipo I, mas também adotaram estratégias específicas para minimizar as reações pró-inflamatórias induzidas por vírus (BANERJEE; BAKER et al., 2020). O interessante é que os estressores ambientais podem reativar o vírus da latência, ocasionando em níveis virais mais elevados em morcegos infectados continuamente. Os níveis virais elevados fornecem um ambiente adequado para a replicação viral, levando a persistência viral, por sua vez, pode ocasionar novas estirpes virais (SUBUDHI; RAPIN; MISRA, 2019).

A partir disso, podemos supor que os morcegos possuem uma capacidade inata adicional de defesa antiviral e, por outro lado, o potencial de andar de mão dadas com vírus para gerar variabilidade. Em consequência, os vírus de morcego são altamente capazes de superar várias barreiras para se tornar um patógeno zoonótico. Os casos de infecção são mais comuns em pessoas do sexo masculino e idosos relacionados a gravidade e morte e, principalmente, aqueles que possuem uma doença subjacente, como diabetes, doenças cardiovasculares e hipertensão (NASAB; SAGHAZADEH; REZAEI, 2020).

O vírus possui uma estrutura esférica e uma glicoproteína inserida no envelope (MAGRONE, MAGRONE, JIRILL, 2020). Essa proteína tem sido demonstrada para se ligar aos receptores celulares do hospedeiro e, assim, facilitar a fusão de membrana (LI, 2016). Sendo que, após o vírus entrar nos pulmões através da respiração, o SARS-CoV-2 estimula a atividade das células imunes, com isso, ativa outros mecanismos de resistência ao patógeno e aumenta a produção de citocinas (ALEGRA et al., 2020).

Os vírus de RNA são identificados pelo sistema imunológico inato através de três grupos de receptores de reconhecimento de padrões, sendo eles receptores semelhantes a RIG-I (RLRs), receptores semelhantes a NOD (NLRs) e receptores semelhantes a Toll (TLRs). Eles estimulam a produção de interferon (IFN) e, dessa forma, desencadeiam efeitos antivirais como células T CD8+, células Natural Killer (NK) e macrófagos (NELEMANS; KIKKERT, 2019; FEHR; CHANNAPPANAVAR; PERLMAN, 2017).

Após a identificação de células infectadas que apresentam os antígenos virais como porções de complexos de antígenos MHC-I da superfície os linfócitos T citotóxicos são estimulados. Assim, a apresentação eficaz é determinada pelo alojamento correto de antígenos por moléculas de MHC-I através de ligações de hidrogênio e relações de ponte salina que permitem grande afinidade com maior especificidade (ROGNAN et al., 1992).

As infecções causadas por SARS-CoV e MERS-CoV são caracterizadas pela rápida replicação inicial do vírus com geração tardia de IFN, resultando em respostas inflamatórias desproporcionais causando alterações pulmonares graves (KINDLER; THIEL, 2016; CHANNAPPANAVAR et al., 2016).

A imunidade dos sujeitos reduz e a virulência do vírus aumenta devido a luta entre o vírus e o corpo humano (DIAO et al., 2020). Com isso, causa congestionamento pulmonar e edema, espessamento do tecido intersticial e exsuação aumentada no espaço alveolar capaz de causar insuficiência respiratória. No funcionamento do sistema imunológico os subconjuntos

de células desempenham um papel central, como as células T CD4+, células T CD8+, células B e células NK (ALLEGRA et al., 2020).

Num estudo, a quantidade total de células T, células B e células NK foi significativamente reduzida no grupo infectado, pois as células T e as células NK estavam abaixo da faixa normal, enquanto as células B estavam dentro da menor quantidade dos valores de referência. As células T são as mais alteradas pela infecção viral, aproximadamente metade do limite de referência inferior. No entanto, a função das células CD4+, CD8+ T e células NK estava dentro dos limites normais deste estudo, como sugerido pelo PMA/Ionicina- $\gamma$  células positivas nessas três populações. Além disso, examinando os vários subconjuntos das células T, os autores evidenciaram que tanto as células-auxiliares (CD3+CD4+) quanto as células T citotóxica (CD3+CD8+) estavam abaixo da faixa normal em indivíduos com COVID-19, com a razão de auxiliar/supressor T (Th/Ts) dentro dos limites normais. Além disso, os sujeitos com infecção pelo SARS-CoV-2 apresentaram menor número de células T regulatórias (Treg) (CD3+CD4+CD25+CD127low+), e essa redução foi especialmente evidente em pacientes críticos (QIN, 2020).

Além do que, esses dados foram confirmados por outros estudos que mostram números de linfócitos abaixo do limite normal na maioria dos indivíduos. As células CD+, CD8+, NK T e linfócitos B foram todas reduzidas em indivíduos infectados. A presença de comorbidades modificou significativamente a quantidade de células T CD8+ (LIU, Z. et al., 2020).

Todos os estudos já concordam em mostrar que as alterações do subconjunto de células T se correlacionam fortemente com a condição inflamatória. A quantidade de células T CD8+ foi negativamente correlacionada, enquanto a razão CD4+/CD8+ correlaciona-se positivamente com proteína C-reativa (PCR), taxa de sedimentação eritrócito (ESR) e IL-6 (ALLEGRA et al., 2020).

Os pacientes infectados apresentaram aumento nas células T CD8 + e células B que responderam positivamente ao tratamento, enquanto não foram relatadas alterações relevantes em qualquer subconjunto de linfócitos em pacientes sem resposta. Assim, linfócitos e seus subconjuntos, principalmente células T CD8+, podem ser um possível marcador para a eficácia clínica dos tratamentos para infecção por SARS-CoV-2 (WANG, F. et al., 2020).

Dessa forma, foram realizados estudos específicos para análise de linfócitos T CD4 +. É relevante saber que a diferenciação das células T CD4+ virgens em efetivas e na população de memória é um dos aspectos mais essenciais da imunidade mediada por células T, e o equilíbrio entre as células T CD4 + virgens e de memória é essencial para manter uma resposta



imune eficaz. Em pacientes críticos (COVID-19) esse equilíbrio é alterado com expansão do subconjunto de células T CD4<sup>+</sup> virgens e redução da porcentagem celular de memória, mostrando ainda mais a gravidade do comprometimento do sistema imunológico nesses indivíduos. Foi observado que a carga de RNA SARS-CoV-2 e a quantidade de linfócitos e a quantidade de linfócitos T CD4<sup>+</sup> e CD8<sup>+</sup> estavam negativamente correlacionados (LIU, Y. et al., 2020).

Assim, esses resultados sugerem que a diminuição dos linfócitos e seus subconjuntos, diretamente afetados pela carga viral, está intimamente relacionada à evolução da doença. Um estudo de citometria de fluxo em relação aos outros subconjuntos de linfócitos demonstrou um aumento no número de células Th17, que se diferenciam das células Th0 estimuladas principalmente por IL-6 e IL-23 (HE et al., 2020).

Para explicar o mecanismo final pelo qual a infecção viral causa uma redução em várias populações de células existem dados interessantes, como a estimulação antigênica. O estudo demonstrou que, lavado broncoalveolar, muitos genes significativamente modificados estão associados à apoptose e às vias de sinalização de P53, compreendendo GTSE1, RRM2, CTSL, CTSB, DDIT4, CCNB1, RRAS, CDK1, CTSD, STEAP3, BIRC5, TNFSF10, CTSZ, NTRK1, IGFBP3, CCNB2 e TP53I3. Um gene essencial no processo de morte celular programada é o TP53. Esse gene exibe uma tendência crescente, sugerindo que a diminuição das células mononucleares do sangue periférico pode ser devido a um aumento na morte celular programada (XIONG et al., 2020).

### **3.3 Análise Filogenética do SARS CoV-2**

Na natureza, um dos mecanismos mais importantes de evolução viral é a substituição de nucleotídeos (LAURING; ANDINO, 2010). O ritmo de mutação do novo Coronavírus como um vírus de RNA é, possivelmente, uma substituição de 10<sup>-4</sup> pb a cada ano. Sendo que, variações podem surgir no curso de cada ciclo de replicação do genoma (SU et al., 2016).

A análise de dados de sequência genética de patógenos é cada vez mais reconhecida como uma ferramenta importante na epidemiologia de doenças infecciosas (RAMBAUT et al., 2008; GRENFELL et al., 2004).

O sequenciamento do genoma do SARS-CoV-2 começou em um estágio inicial do surto em Wuhan. As amostras do fluido de lavagem broncoalveolar foram coletadas dos pacientes iniciais. Este grande volume de dados de sequenciamento fornece uma grande oportunidade para identificar as variações de estirpes do vírus, a tendência de evolução /

mutações do vírus e seu impacto no potencial patogênico do SARS-CoV-2 (KADAM et al., 2021).

Assim, a sequência do genoma do novo vírus agora está disponível em diferentes partes do mundo. Várias mutações no genoma viral foram identificadas usando essas sequências de genoma completos. Ainda não existem correlações concretas derivadas entre as mutações relatadas e a patogenicidade / propagação do vírus. Mas, o estudo dessas mutações será definitivamente útil para prever a futura trajetória de surtos de vírus e também para o desenvolvimento de tratamentos antivirais e vacinas (LI, X. et al., 2020).

Os vírus são propensos a mutações rápidas, o que os permite escapar da imunidade do hospedeiro (STERN; ANDINO, 2016). A proteína spike desse vírus é a principal proteína viral que contribui para a fusão da partícula do vírus com a célula hospedeira por meio da enzima conversora de angiotensina 2 (ECA2). A expressão altamente conservada de ECA2 é encontrada em vários animais, o que indica sua função fisiológica central. Várias variantes dessa enzima podem afetar a estabilidade da proteína, influenciando a interação entre a proteína spike e ECA2 por meio da imposição de mudanças conformacionais, enquanto algumas outras variantes são conhecidas por causar uma diminuição ou um aumento na afinidade ligante-receptor (BAKHSHANDEH et al., 2021). A ECA2 é encontrada na superfície das células dos rins, do endotélio, dos pulmões e, também, em outros órgãos (SCHOLZ et al., 2020).

Dessa forma, acredita-se que a interação entre as proteínas da spike coronaviral e seu receptor de fixação ECA2 seja a chave para a transmissibilidade da infecção para órgãos e tecidos (GU; KORTEWEG, 2007).

Além disso, variantes específicas nos genes ECA2 e TMPRSS2 foram relatadas entre diversas populações em todo o mundo, sugerindo que diferentes origens genéticas do hospedeiro podem contribuir para as diferenças na infecção e na gravidade do COVID-19. A propagação da infecção por essa doença apresenta variação geográfica. A maioria das mortes relacionadas a essa doença está sendo relatada nas Américas, Europa e Sudeste Asiático, com menos mortes na Oceania e no Leste Asiático (SECOLIN et al., 2021).

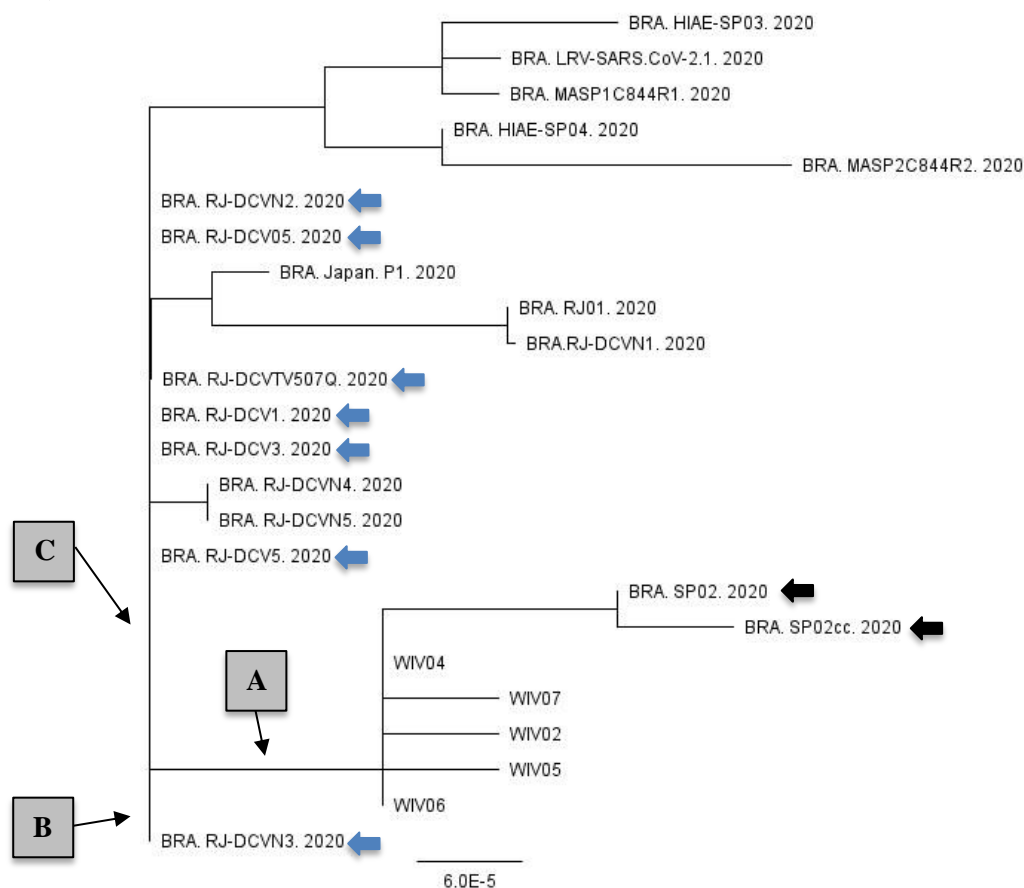
Dessa forma, o acesso às sequências genômicas da população em geral é relevante para a identificação da variabilidade genética envolvida nos mecanismos moleculares de infecção (CHAPMAN; HILL, 2012).

Neste trabalho, foram utilizados 105 genomas completos para a criação das árvores filogenéticas, incluindo os países: Brasil, China, Índia, Itália e Estados Unidos. E, além disso, 5 sequências WIV. O alinhamento das sequências foi realizado pelo programa Geneious. Os

nomes das sequências estão de acordo com a publicação do genoma completo na plataforma do Genbank. E uma árvore filogenética é uma representação gráfica que representa relações evolutivas entre organismos que possam ter um ancestral comum. Sendo que, é chamando de clado ou ramo cada ramificação da árvore que significa que os organismos correspondem de um ancestral comum.

Em relação a COVID-19, o Brasil relatou o terceiro maior número de casos em todo o mundo (atualizado em 09 de julho de 2021). Sendo que, o maior número de casos concentra-se na região sudeste do país (atualizado em 28 de setembro de 2020) (MOURA et al., 2015; KEHDY et al., 2015; LIMA-COSTA et al., 2015). Dessa forma, pôde ser observado o maior número de publicações da região sudeste na plataforma Genbank. Na figura 4, é possível observar a organização das sequências de acordo com sua similaridade genética.

**Figura 4:** Análise filogenética dos principais genomas virais do Coronavírus encontrados em hospedeiros humanos no Brasil comparados com genomas de Coronavírus com provável origem de (WIV).



Fonte: Elaborada pela autora.

De acordo com a figura 4, apenas duas sequências tiveram semelhança com as sequências WIV, sendo a BRA. SP02.2020 e BRA. SP02cc.2020 (sinalizadas pela seta preta).

Com isso, demonstra que as estirpes brasileiras da pesquisa possuem uma alta variabilidade genética que pode ser observado pelas sequências BRA. RJ-DCVN2.2020, BRA. RJ-DCV05.2020, BRA. RJ-DCVTV507Q. 2020, BRA. RJ-DCV1. 2020, BRA. RJ-DCV3. 2020, BRA. RJ-DCV5. 2020, BRA. RJ-DCVN3. 2020 (indicadas pela seta azul) que não se encontram em ramificação com outras estirpes.

**Quadro 1:** Representativo da distribuição das estirpes do Brasil com os genomas de Coronavírus com provável origem de morcegos (WIV).

<b>Grupamentos</b>	<b>Estirpes</b>
<b>A</b>	Possui todas as sequências WIV e apenas duas estirpes estão relacionadas (BRA. SP02. 2020 e BRA. SP02cc. 2020).
<b>B</b>	Apenas um ramo com uma única estirpe (BRA. RJ-DCVN3. 2020).
<b>C</b>	É possível observar ramificações e subgrupos a partir da ramificação principal (C).

Fonte: Elaborado pela autora.

Uma característica dos indivíduos brasileiros é um genoma misto, abrangendo europeus, africanos subsaarianos e nativos americanos, e a distribuição dos componentes ancestrais varia notavelmente ao longo do genoma (MOURA et al., 2015; KEHDY et al., 2015; LIMA-COSTA et al., 2015).

Além disso, sabe-se que a população miscigenada brasileira é sub-representada em grandes bancos de dados públicos (KARCZEWSKI et al., 2020; AUTON et al., 2015) e estudos anteriores revelaram variantes presentes exclusivamente em indivíduos brasileiros (ROCHA et al., 2020; NASLAVSKY et al., 2017). Assim, supostamente que variantes raras ou ultrararas de grande efeito podem dar uma contribuição essencial para a infecção por COVID-19 nessa população (SECOLIN et al., 2021). Isso é confirmado na observação da árvore, devido muitas estirpes não apresentarem aproximação evolutiva, assim, demonstrando que houve mutações que os afastam.

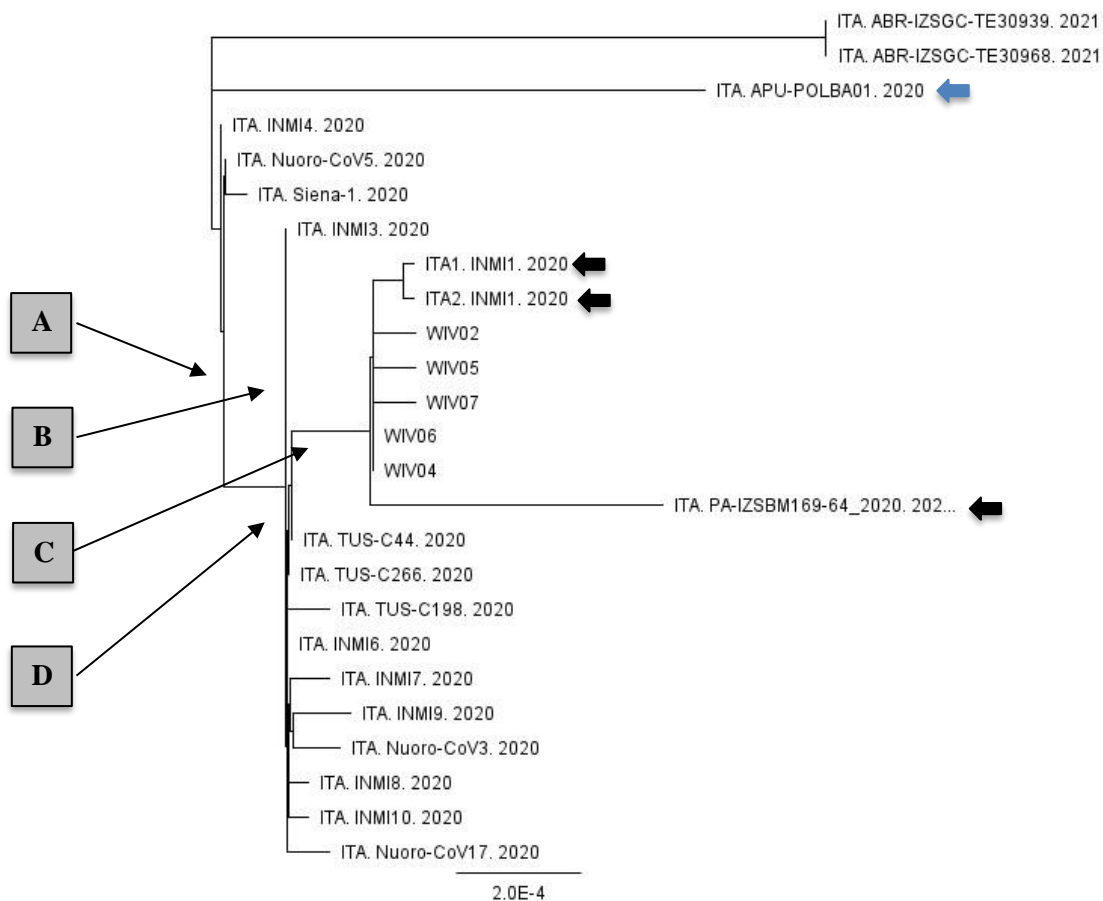
Além disso, foi demonstrado que uma proporção significativa da variabilidade genética em brasileiros miscigenados permanece desconhecida (ROCHA et al., 2020) e que a variabilidade genética pode levar a uma resposta diferencial à infecção (CHAPMAN; HILL, 2012).

Estudos anteriores mostraram que os genes ECA2, TMPRSS2, CTSL, FURIN e IL6, bem como a região HLA, podem estar envolvidos na infecção por SARS-CoV-2 e na resposta

imune. Além disso, as variantes nos loci 3p21.31 e 9q34.2 (abrangendo SLC6A20, LZTFL1, FYCO1, CXCR6, XCR1, CCR9 e ABO) foram associadas a pacientes espanhóis e italianos com COVID-19, e diferentes variantes afetam a predisposição para a vida (ZHANG et al., 2020).

Na Itália, é possível observar, na figura 5, as ramificações que foram formadas englobando todas as sequências selecionadas para o estudo. Apenas 3 sequências estão relacionadas com as sequências WIV: ITA1. INMI. 2020, ITA2. INMI1. 2020 e ITA. PA-IZSBM169-64\_2020. 202 (indicadas pela seta preta). Diferente do Brasil, a Itália possui uma similaridade maior entre suas sequências, devido todos os genomas estarem em ramificação com outras estirpes.

**Figura 5:** Análise filogenética dos principais genomas virais do Coronavírus encontrados em hospedeiros humanos na Itália comparados com genomas de Coronavírus com provável origem de morcegos (WIV).



Fonte: Elaborada pela autora.

As diferentes taxas de mortalidade e infecção de COVID-19 em áreas geográficas podem ser influenciadas por vários fatores, como diferentes regras nacionais de quarentena e

restrições de movimento, e também diferentes raças ou antecedentes genéticos das pessoas (KAWAGUCHI et al., 2018). A análise de dados de matriz SNP revelou que um haplótipo específico do gene TMPRSS2 em uma população europeia que resulta na regulação positiva da expressão do gene TMPRSS2 (como uma enzima crucial para a entrada de SARS-CoV-2 na célula hospedeira) de uma maneira dependente de andrógeno, mostrou associação com a gravidade da doença na população italiana (HALACHEV et al., 2019).

Apesar de possuírem silimaridade de forma geral devido as ramificações, é possível também compreender suas diferenças ao analisar cada clado da árvore acima, como a ITA. APU-POLBA01. 2020 (indicado pela seta azul). Essa estirpe se apresenta sozinha, logo mutações em seu genoma foram observados durante o processo de alinhamento em relação as demais. Além disso, as mutações virais ao longo do tempo são consideradas os fatores que contribuem para a transmissão e gravidade da infecção por COVID-19 (KAWAGUCHI; HIGASA et al., 2018). No quadro 2, está descrito as características principais da árvore filogenética da Itália.

**Quadro 2:** Representativo da distribuição das estirpes da Itália com os genomas de Coronavírus com provável origem de morcegos (WIV).

<b>Grupamentos</b>	<b>Estirpes</b>
<b>A</b>	Apresenta subgrupos a partir da ramificação principal (A).
<b>B</b>	Único clado representado pela estirpe ITA. INMI3. 2020.
<b>C</b>	Apresenta ramificação com as estirpes (ITA1. INMI. 2020, ITA2. INMI1. 2020 e ITA. PA-IZSBM169-64_202...) relacionadas com as sequências WIV.
<b>D</b>	Grupamento com maior número de estirpes relacionadas e apresenta subgrupos a partir da ramificação principal (D).

Fonte: Elaborado pela autora.

Enquanto o genoma SARS-CoV-2 é mais estável do que SARS-CoV ou MERS-CoV, ele tem uma taxa de mutação dinâmica relativamente alta comparável a outros vírus de RNA. Algumas mutações podem ser como mutações fundadoras e mostrar padrões geográficos específicos. Por exemplo, a mutação D614G na proteína S é uma mutação fundadora na Europa (aparecimento em 16 de fevereiro de 2020) (SECOLIN et al., 2021).

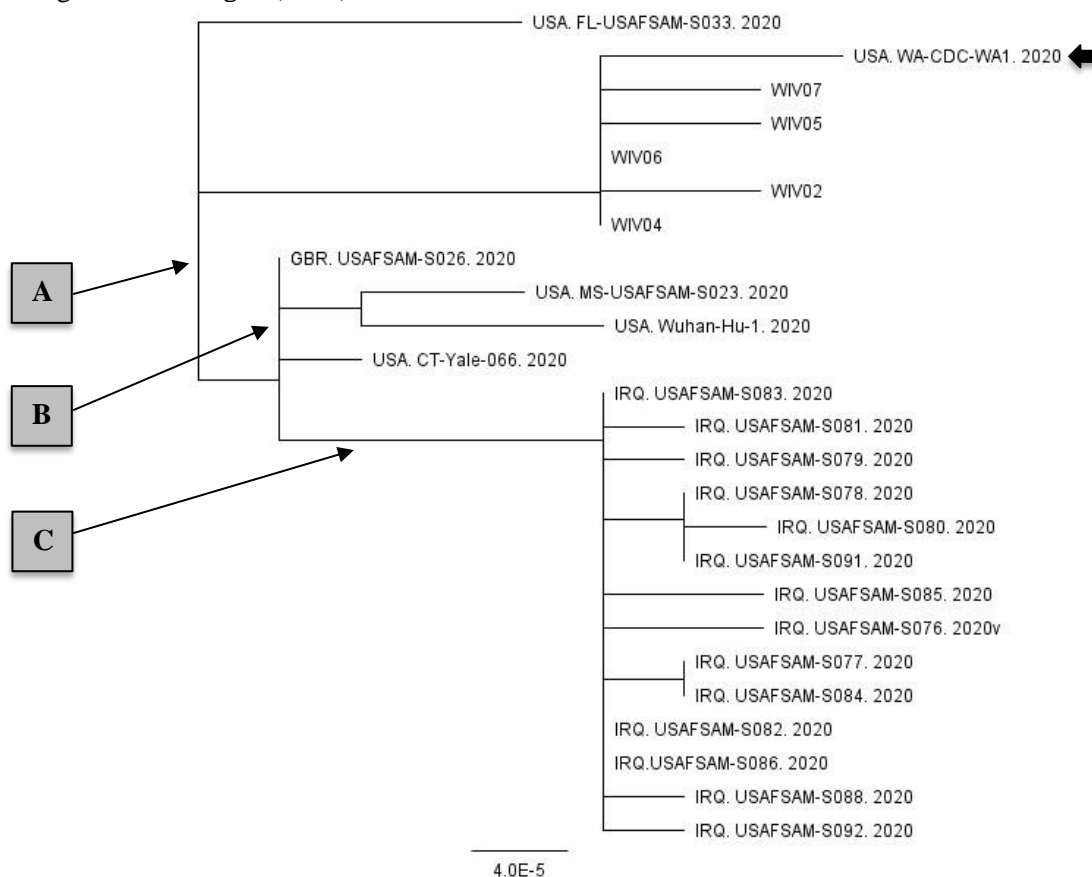
Sem dúvida, essas mutações podem conduzir a variabilidade genética viral, e por causa da correlação genótipo-fenótipo, resultar em um vírus com mais, menos ou nenhuma

diminuição na aptidão patogênica natural ou, por outro lado, facilitando sua rápida mudança antigênica para escapar do sistema imunológico do hospedeiro e, também, criando um vírus de resistência aos medicamentos (SCHWARZ et al., 2010; GONZÁLEZ-PÉREZ; LÓPEZ-BIGAS, 2011).

Por conseguinte, a variação do genoma do vírus detectada em pacientes europeus e norte-americanos (principalmente na Europa) mostrou um número maior do que na Ásia durante a propagação do vírus para fora da China e, como resultado, a infecção tem uma taxa de mortalidade mais alta fora da China (KAWAGUCHI; HIGASA et al., 2018).

Nos Estados Unidos, temos um perfil um pouco diferente em relação a árvore da Itália, conforme a figura 6. Apenas a sequência USA. WA-CDC-WA1. 2020 (indicada pela seta preta) tem similaridade com as sequências WIV. Sendo elas as primeiras sequências publicadas em relação a pneumonia por SARS-CoV-2 com possível origem de morcego. Após a figura 6, o quadro 3 expõe as características da árvore filogenética.

**Figura 6:** Análise filogenética dos principais genomas virais do Coronavírus encontrados em hospedeiros humanos nos Estados Unidos comparados com genomas de Coronavírus com provável origem de morcegos (WIV).



Fonte: Elaborada pela autora.

**Quadro 3:** Representativo da distribuição das estirpes dos Estados Unidos com os genomas de Coronavírus com provável origem de morcegos (WIV).

<b>Grupamentos</b>	<b>Estirpes</b>
<b>A</b>	Grupamento que possui uma estirpe (USA. WA-CDC-WA1. 2020) relacionada com as WIV. Possui subgrupos a partir da ramificação principal (A).
<b>B</b>	Apresenta subgrupos a partir da ramificação principal (B).
<b>C</b>	Apresenta o maior número de sequências relacionadas em uma ramificação com presença de subgrupos a partir da ramificação principal (C).

Fonte: Elaborado pela autora.

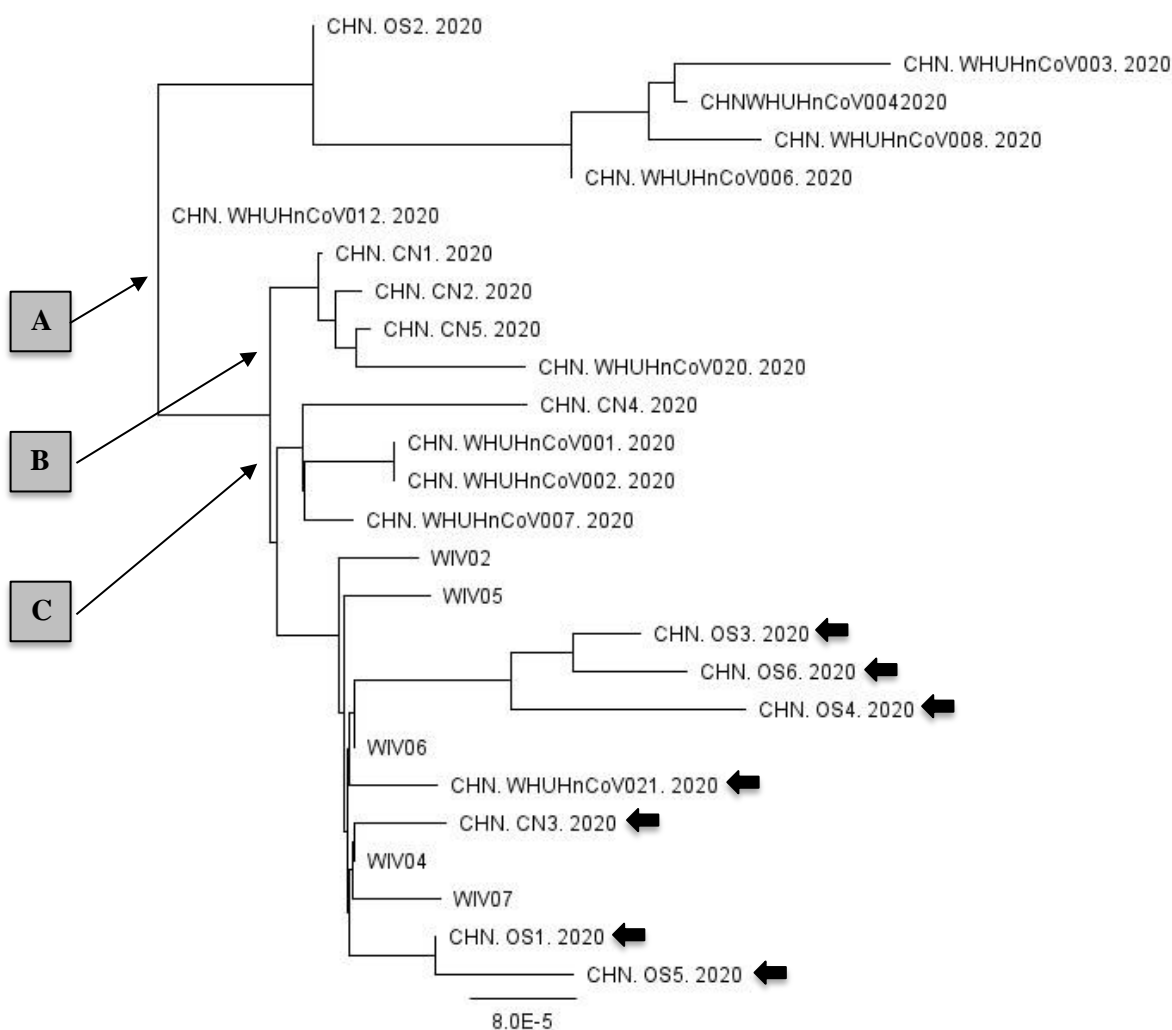
Assim, os agentes causadores das mutações do genoma do vírus são vários, incluindo fidelidade de RdRp (como uma enzima de replicação / transcrição principal do vírus), enzimas do hospedeiro, efeitos de mutagênicos físicos e químicos e recombinação genética. Além disso, após a ocorrência da mutação RdRp 14408 nt (a partir de 16 de fevereiro de 2020), o padrão de outras variações foi alterado e novas versões foram introduzidas no genoma do vírus na Europa. No entanto, o padrão de mutação semelhante na América do Norte não foi observado nas sequências do genoma SARS-CoV-2 isoladas de pessoas infectadas, que tinham mutação RdRp 14408 nt. Como se, a ocorrência simultânea de mutação RdRp 14408 nt com nsp3 3036 nt, proteína S 23403 nt e fosfoproteína de nucleocapsídeo 28881 nt na Europa e com nsp13 (helicase) 17746 nt e 17857 nt, e nsp14 (ExoN) 18060 nt na América do Norte tinha acontecido (KAWAGUCHI; HIGASA et al., 2018).

Já na China, como mostra a figura 7, é o país com maior número de sequências relacionadas com as WIV, sendo elas CHN. OS3. 2020, CHN. OS6. 2020, CHN. OS4. 2020, CHN. WHUHnCoV021. 2020, CHN. CN3. 2020, CHN. OS1. 2020 e CHN. OS5. 2020 (sinalizadas pela seta preta). Em seguida, após a figura, o quadro 4 apresenta as características observadas na árvore filogenética da China.

Sendo que, o surto de COVID-19 na China atingiu o pico epidêmico em fevereiro. De acordo com a Comissão Nacional de Saúde da China, o número total de casos continuou a aumentar acentuadamente no início de fevereiro, a uma taxa média de mais de 3.000 novos casos confirmados por dia. Para controlar a COVID-19, a China implementou medidas de saúde pública estritas e sem precedentes (HU et al., 2020).



**Figura 7:** Análise filogenética dos principais genomas virais do Coronavírus encontrados em hospedeiros humanos na China comparados com genomas de Coronavírus com provável origem de morcegos (WIV).



Fonte: Elaborada pela autora.

**Quadro 4:** Representativo da distribuição das estirpes da China com os genomas de Coronavírus com provável origem de morcegos (WIV).

Grupamentos	Estirpes
<b>A</b>	Apresenta subgrupos a partir da ramificação principal (A). Único grupamento com uma estirpe isolada (CHN. WHUHnCoV12. 2020).
<b>B</b>	Grupamento com ramificações relacionadas entre si.
<b>C</b>	Apresenta grupamento com maior número de ramificações e subgrupos a partir da ramificação principal (C). E, também, o país com o maior número de estirpes relacionadas com as seqüências WIV.

Fonte: Elaborado pela autora.

A cidade de Wuhan foi fechada em 23 de janeiro e todas as viagens e transportes que ligavam a cidade foram bloqueados. Nas duas semanas seguintes, todas as atividades ao ar livre e reuniões foram restritas e as instalações públicas foram fechadas na maioria das cidades, bem como no campo (SECOLIN et al., 2019). Devido a essas medidas, o número diário de novos casos na China começou a diminuir continuamente (ROCHA; SECOLIN et al., 2020).

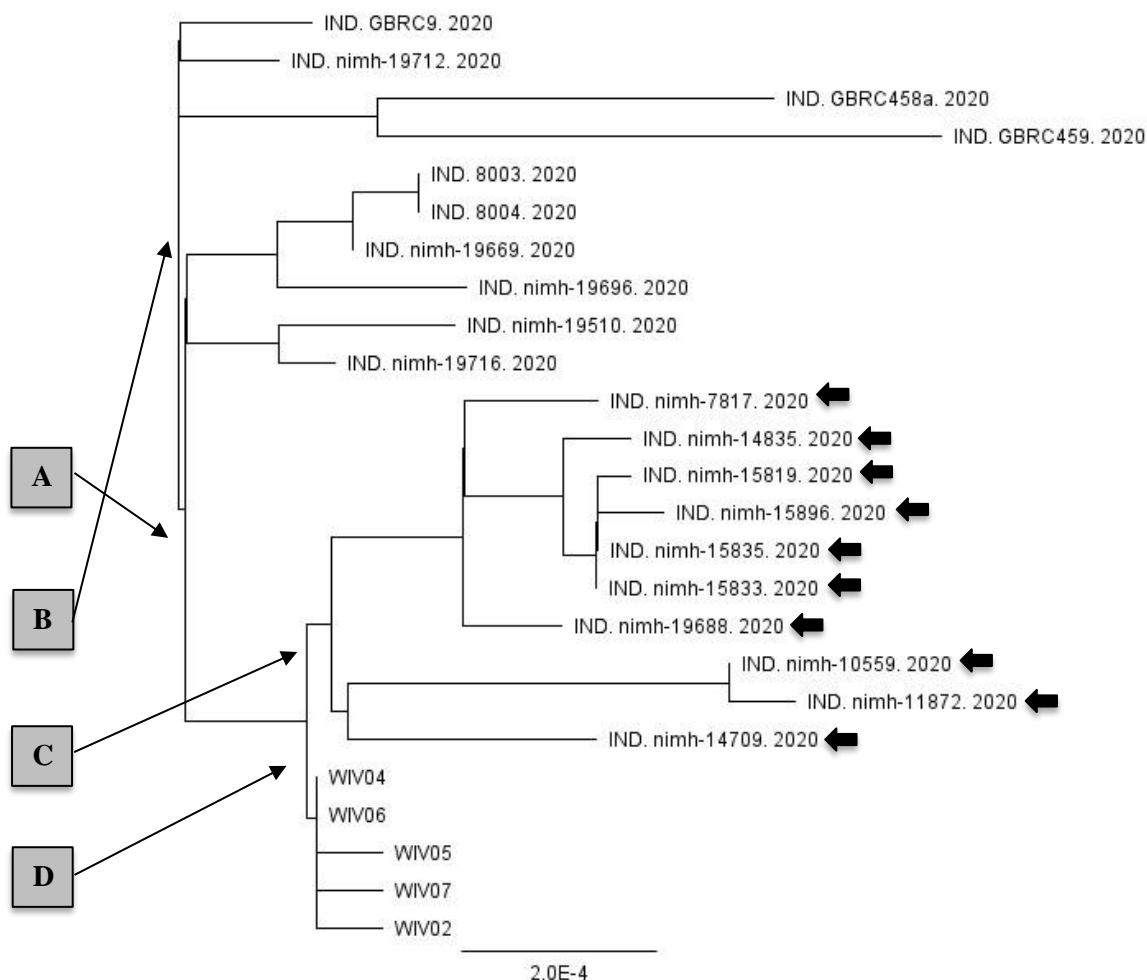
No entanto, apesar da tendência de queda na China, grandes grupos de infecção têm sido relatados em um número crescente de países (SECOLIN et al., 2019). A alta eficiência de transmissão do SARS-CoV-2 e a abundância de viagens internacionais possibilitaram a rápida disseminação do COVID-19 em todo o mundo. Em 11 de março de 2020, a OMS caracterizou oficialmente o surto global de COVID-19 como uma pandemia (CHAPMAN; HILL, 2012).

Desde março, enquanto o COVID-19 na China passou a ser efetivamente controlado, o número de casos na Europa, nos Estados Unidos e em outras regiões aumentou drasticamente. De acordo com o painel COVID-19 do Center for System Science and Engineering da Johns Hopkins University (Centro de Ciência e Engenharia de Sistemas da Universidade Johns Hopkins), em 11 de agosto de 2020, 216 países e regiões de todos os seis continentes relataram mais de 20 milhões de casos de COVID-19 e mais de 733.000 pacientes morreram (KARCZEWSKI et al., 2020). A alta mortalidade ocorreu especialmente quando os recursos de saúde estavam sobrecarregados. Os Estados Unidos são o país com maior número de casos até agora (HU et al., 2020).

Além disso, a filogenia das sequências do genoma SARS-CoV-2 mostra uma estruturação geográfica que surgiu devido a evoluções locais após a dispersão geograficamente de longa distância do vírus (JONES; MANRIQUE, 2020). Por exemplo, 30% da sequência proteômica de Wuhan-Hu-1 não mostra nenhuma semelhança com a de isolados indianos (PRATHIVIRAJ; KIRAN; SELVIN, 2020).

Na Índia, podemos observar, na figura 8, uma relação de afinidade por um ancestral comum entre algumas sequências e as sequências WIV. Porém, não possuem uma similaridade direta com as sequências WIV. São elas: IND. nimh-7817. 2020, IND. nimh-14835. 2020, IND. nimh-15819. 2020, IND. nimh-15896. 2020, IND. nimh-15835. 2020, IND. nimh-15833. 2020, IND. nimh-19688. 2020, IND. nimh-10559. 2020, IND. nimh-11872. 2020 e IND. nimh-14709. 2020 (indicadas pela seta preta). Logo após a figura, o quadro 5 apresenta o perfil da árvore filogenética da Índia.

**Figura 7:** Análise filogenética dos principais genomas virais do Coronavírus encontrados em hospedeiros humanos na Índia comparados com genomas de Coronavírus com provável origem de morcegos (WIV).



Fonte: Elaborada pela autora.

**Quadro 5:** Representativo da distribuição das estirpes da Índia com os genomas de Coronavírus com provável origem de morcegos (WIV).

Grupamentos	Estirpes
<b>A</b>	Apresenta subgrupos a partir da ramificação principal (A).
<b>B</b>	Apresenta ramificações e subgrupos a partir da ramificação principal (B).
<b>C</b>	Grupamento com o maior número de estirpes. Apresenta, também, subgrupos a partir da ramificação principal (C).
<b>D</b>	Não possui estirpes relacionadas diretamente com as sequências WIV.

Fonte: Elaborado pela autora.

É possível observar que, diferente dos demais países selecionados para o estudo, a Índia possui um perfil de ramificação diferenciado. Dessa forma, contém mais ramificações isoladas. Estudos revelaram que a Europa e o Sudeste Asiático são as duas principais rotas de introdução da doença na Índia, seguida de transmissão local (RAGHAV et al., 2020).

Geralmente, a mutação ocorre quando um erro é incorporado em um genoma viral. Também pode ser considerado um mecanismo de enfrentamento de danos genômicos. Como consequência, a estirpe mutada resultante pode causar um surto em hospedeiros humanos, como no caso do SARS-CoV-2 (SAHA et al., 2020).

Entre todas as proteínas acessórias desse vírus, a ORF3a é considerada muito importante na patogênese do vírus, uma vez que possui várias mutações que estão ligadas às proteínas spike. Assim, existem diferentes mutações que ocorrem em vários locais do gene ORF3a dos genomas SARS-CoV-2 de pacientes indianos e essas mutações levam à alternância de aminoácidos (HASSAN et al., 2020).

O monitoramento contínuo e a análise das sequências dos países afetados seriam vitais para entender a evolução genética e as taxas de substituição do SARS-CoV-2. Essa diversidade genética das populações desse vírus circulando em diferentes países indica que cada uma dessas epidemias locais foi semeada por um grande número de introduções independentes do vírus. A principal exceção a esse padrão é a China, a fonte do surto inicial, onde apenas uma fração da diversidade global está presente. Este também é o caso da Itália, que foi um dos primeiros focos da pandemia COVID-19.

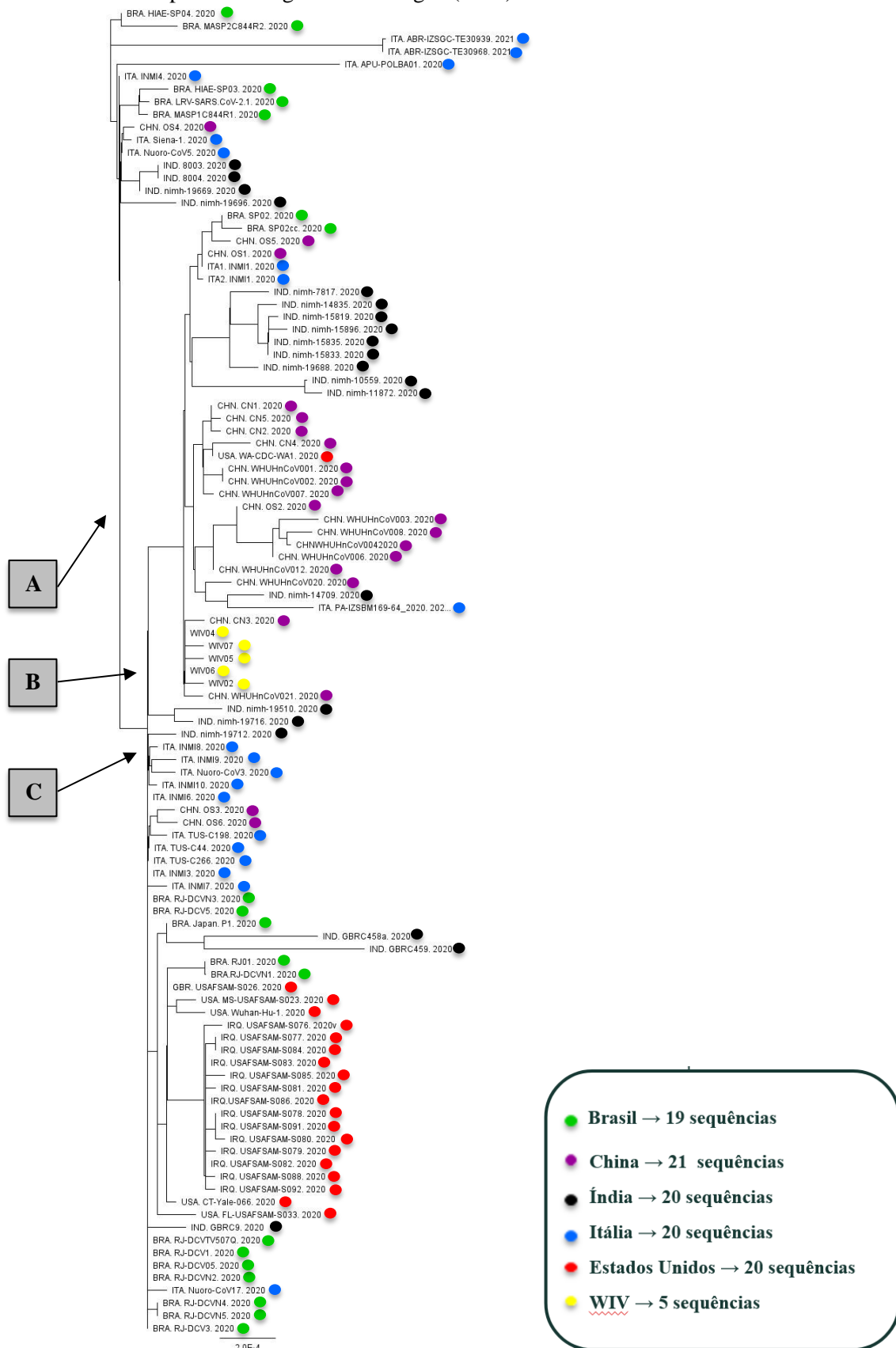
Na figura 8, é possível visualizar de uma forma geral a organização das sequências dos países e sua intereção em relação a proximidade que as estirpes possuem entre si. Com a cor verde (Brasil), vermelho (Estados Unidos), azul (Itália), preto (Índia), roxo (China) e amarelo (sequências WIV). A diversidade genômica da população global de SARS-CoV-2 sendo recapitulada em vários países aponta para uma extensa transmissão mundial de COVID-19 (YADAV et al., 2020). Após a figura, o quadro 6 apresenta as características filogenéticas principais.

A filogenia revelou heterogeneidade emergente dentro das sequências SARS-CoV-2 globalmente. Os vírus dos Estados Unidos, da China e da Índia foram posicionados, em sua maioria, reunidos com as estirpes do próprio país. Dessa forma, demonstrando através do alinhamento que mutações diferentes ocorrem pelo mundo. Além disso, percebe-se que algumas estirpes tem proximidade com outras sequências de outros países devido a circulação de pessoas entre as nacionalidades estudadas neste trabalho. Assim, quando existe similaridade

em relação a um ancestral comum elas se apresentam próximas da mesma ramificação, como a USA. WA-CDC-WA1. 2020.

A medida que as infecções por SARS-CoV-2 encontram populações suscetíveis em todo o mundo, e à medida que a taxa de replicação viral aumenta nessas populações, o número de novos pontos de mutação e as frequências de mutações em pontos de acesso individuais aumentaram. Permitir que o vírus se replique com eficiência crescente pode tornar seu genoma mais bem adaptado e cada vez mais perigoso para a saúde humana (WEBER; RAMIREZ; DOERFLER, 2020).

**Figura 8:** Análise filogenética dos principais genomas virais do Coronavírus encontrados em hospedeiros humanos no Brasil, Itália, Estados Unidos, China e Índia comparados com com genomas de Coronavírus com provável origem de morcegos (WIV).



Fonte: Elaborada pela autora.

**Quadro 6:** Representativo da distribuição das estirpes dos principais grupamentos de Coronavírus da árvore filogenética com os 105 genomas completos (A, B e C).

<b>Grupamentos</b>	<b>Estirpes</b>
<b>A</b>	<p>Possui um padrão heterogêneo incluindo quatro países, com exceção dos Estados Unidos. E, também, possui apenas uma estirpe chinesa relacionada (representada pela cor roxa).</p> <p>Grupamento com menos ramificações.</p> <p>É possível observar a partir da ramificação principal (A) a presença de subgrupos que dão origem a outras ramificações de acordo com a sua aproximação genética em relação a um ancestral comum.</p>
<b>B</b>	<p>Possui um padrão heterogêneo e inclui todos os países, inclusive as sequências WIV (isoladas a partir de humanos).</p> <p>Grupamento relacionado com o maior número de estirpes chinesas.</p> <p>É possível observar a partir da ramificação principal (B) a presença de subgrupos que dão origem a outras ramificações de acordo com a sua aproximação genética em relação a um ancestral comum.</p>
<b>C</b>	<p>Possui um padrão heterogêneo e inclui todos os países. É o maior grupamento e possui relação com somente duas estirpes chinesas. Além disso, apresenta o maior número de estirpes de um mesmo país (Estados Unidos) numa mesma ramificação.</p> <p>Grupamento com o maior número de ramificações.</p> <p>É possível observar a partir da ramificação principal (C) a presença de subgrupos que dão origem a outras ramificações de acordo com a sua aproximação genética em relação a um ancestral comum.</p>

Fonte: Elaborado pela autora.

#### 4. CONCLUSÃO

De acordo com o exposto, foi evidenciado a capacidade infectiva e disseminativa do SARS-CoV-2. Além disso, a suscetibilidade de grupos populacionais bem como os processos de replicação viral que, a partir disso, acionam mecanismos genéticos adaptativos que tornam o vírus potencialmente proliferativo dentro das células, com isso, provocando sinais e sintomas. Assim, fica dependente das condições físicas e biológicas do hospedeiro o que pode resultar em uma evolução do quadro levando o indivíduo a óbito.

Com base nos dados apresentados, é possível afirmar que esse novo Coronavírus é originado de um processo de coevolução (interação genética entre vírus e hospedeiro) de uma

extensa linhagem que acredita-se ter início em origem animal e atingiu os hospedeiros humanos através do salto de espécies. Dessa forma, com base na filogenia molecular do genoma viral essa capacidade zoonótica traz consigo o alerta para possíveis novos surtos ou epidemias devido à evolução e adaptações virais em hospedeiros animais.

Em geral, os vírus de RNA codificam proteínas que podem se proliferar alterando características celulares e funcionais. Assim, através da análise detalhada de bases moleculares, biológicas e filogenéticas é possível observar que grande parte dos vírus da família *Coronaviridae* utilizam hospedeiros animais como reservatório para sua evolução e desenvolvimento. A partir disso, permite a recombinação de genomas até mesmo entre estirpes diferentes.

De fato, a pandemia que se instaurou em todo o mundo utiliza-se de mecanismos genéticos complexos e similares a eventos já ocorridos em outros países, como a epidemia de síndrome respiratória aguda grave (SARS) e a síndrome respiratória do Oriente Médio (MERS). Mas, através do mecanismo evolutivo e adaptativo o SARS-CoV-2 se espalhou com facilidade chegando a todos os continentes, com ambientes adversos e populações variadas, o que torna ainda mais essencial o monitoramento das estirpes e novas variantes com potencial ainda mais infectivo e virulento.

Por isso, é necessário reafirmar as ações preventivas no combate ao vírus, desde as restrições ao contato entre as pessoas (distanciamento social) como, também, a importância do isolamento social e a vacinação em massa. A vacinação, por exemplo, reduz drasticamente os números de casos graves e contaminados por população e, com isso, diminuem as taxas de mortalidade e morbidade da doença. São medidas que buscam evitar a propagação e dispersão viral, recombinações e, dessa forma, o surgimento de novas variantes.

## REFERÊNCIAS

ALLEGRA, A. et al. Immunopathology of SARS-CoV-2 Infection: Immune Cells and Mediators, Prognostic Factors, and Immune-Therapeutic Implications. **International journal of molecular sciences**, Basel, v. 21, n. 13, p. 4782, jul. 2020. doi: 10.3390/ijms21134782.

AL-QAANEH, A. M. Genome composition and genetic characterization of SARS-CoV-2. **Saudi journal of biological sciences**, Riyadh, v. 28, n. 3, p. 1978–1989, mar. 2021. doi: 10.1016/j.sjbs.2020.12.053.

ANDERSEN, K. G. et al. The proximal origin of SARS-CoV-2. **Nature medicine**, New York, v. 17, p. 1-3, mar. 2020. doi: 10.1038/s41591-020-0820-9.



AUTON, A. et al. A global reference for human genetic variation. **Nature**, London, v. 526, n. 7571, p. 68-74, out. 2015. doi: 10.1038/nature15393.

AZEEZ, S. A. State-of-the-art tools to identify druggable protein ligand of SARS-CoV-2. **Archives of Medical Science: AMS**, Poland, v. 16, n. 3, p. 497–507, mar. 2020. doi: 10.5114/aoms.2020.94046.

BAKHSHANDEH, B. et al. Variants in ACE2; potential influences on virus infection and COVID-19 severity. **Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases**, Amsterdam, v. 90, p. 104773, fev. 2021. doi: 10.1016/j.meegid.2021.104773.

BANERJEE, A.; BAKER M. L. et al. Novel Insights Into Immune Systems of Bats. **Frontiers in immunology**, Lausanne, v. 11, p. 26, jan. 2020. doi: 10.3389/fimmu.2020.00026.

BANERJEE, A. K. et al. SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. **Cell**, Cambridge, v. 183, n. 5, p. 1325-1339.e21, nov. 2020. doi: 10.1016/j.cell.2020.10.004.

BENVENUTO, D. The 2019-new coronavirus epidemic: Evidence for virus evolution. **Journal of medical virology**, New York, v. 92, n. 4, p. 455-459, fev. 2020. doi: 10.1002/jmv.25688.

CERAOLO, C.; GIORGI, F. M. Genomic variance of the 2019-nCoV coronavirus. **Journal of medical virology**, New York, v. 92, n. 5, p. 522-528, mai. 2020. doi: 10.1002/jmv.25700.

CHAN, J. F. et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. **Emerging microbes & infections**, New York, v. 9, n. 1, p. 221–236, jan. 2020. doi: 10.1080/22221751.2020.1719902.

CHANNAPPANAVAR, R. et al. Dysregulated Type I Interferon and Inflammatory Monocyte-Macrophage Responses Cause Lethal Pneumonia in SARS-CoV-Infected Mice. **Cell host & microbe**, Cambridge, v. 19, n. 2, p. 181-193, fev. 2016. doi: 10.1016/j.chom.2016.01.007.

CHAPMAN, S. J.; HILL A. V. S. Human genetic susceptibility to infectious disease. **Nature reviews. Genetics**, London, v. 13, n. 3, p. 175-88, fev. 2012. doi: 10.1038/nrg3114.

CHEN, Y. et al. Emerging coronaviruses: Genome structure, replication, and pathogenesis. **Journal of medical virology**, New York, v. 92, n. 4, p. 418–423, abr. 2020. doi: 10.1002/jmv.25681.

CHEN, J.; LI F.; SHI, Z. Origin and evolution of pathogenic coronaviruses. **Nature reviews. Microbiology**, London, v. 17, n. 3, p. 181-192, mar. 2019. doi: 10.1038/s41579-018-0118-9.

DIAO, B. et al. Reduction and Functional Exhaustion of T Cells in Patients With Coronavirus Disease 2019 (COVID-19). **Frontiers in immunology**, Lausanne, v. 11, p. 827, mai. 2020. doi: 10.3389/fimmu.2020.00827.

DROSTEN, C. et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. **The new england journal of medicine**, Boston, v. 348, n. 20, p. 1967-76, mar. 2003. doi: 10.1056/NEJMoa030747.

DUFFLY, S. Why are RNA virus mutation rates so damn high?. **PLoS biology**, San Francisco, v. 16, n. 8, ago. 2018. doi: 10.1371/journal.pbio.3000003.

ECKERLE, L. D. et al. Infidelity of SARS-CoV Nsp14-Exonuclease Mutant Virus Replication Is Revealed by Complete Genome Sequencing. **PLoS pathogens**, v. 6, n. 5, p. e1000896, mai. 2010. doi: 10.1371/journal.ppat.1000896.

FEHR, A. R.; CHANNAPANNAVAR, R.; PERLMAN S. Middle East respiratory syndrome (MERS): Emergence of a pathogenic human Coronavirus. **Annual review of medicine**, Palo Alto, v. 68, p. 387-399, ago. 2017. doi: 10.1146/annurev-med-051215-031152.

FORSTER, P. et al. Phylogenetic network analysis of SARS-CoV-2 genomes. **Proceedings of the national academy of sciences of the united states of america**, Washington, v. 117, n. 17, p. 9241-9243, abri. 2020. doi: 10.1073/pnas.2004999117.

GAO, H. et al. From SARS to MERS: evidence and speculation. **Frontiers of medicine**, Beijing, v. 10, n. 4, p. 377-382, dez. 2016. doi: 10.1007/s11684-016-0466-7.

GAO, Y. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. **Science**, New York, v. 368, n. 6492, p. 779-782, mai. 2020. doi: 10.1126/science.abb7498.

GE, X. et al. Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. **Virologica sinica**, Beijing, v. 31, n. 1, p. 31-40, fev. 2016. doi: 10.1007/s12250-016-3713-9.

GONZÁLEZ-PÉREZ, A.; LÓPEZ-BIGA, S. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condal. **American journal of human genetics**, Baltimore, v. 88, n. 4, p. 440-9, abr. 2011. doi: 10.1016/j.ajhg.2011.03.004.

GORBALENYA, A. et al. Virus-encoded proteinases and proteolytic processing in the Nidovirales. **The journal of general virology**, London, v. 81, n. Pt 4, p. 853-79, abr. 2000. doi: 10.1099/0022-1317-81-4-853.

GORDON, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. **Nature**, London, v. 583, n. 7816, p. 459-468, jul. 2020. doi: 10.1038/s41586-020-2286-9.

GRENFELL, B. T. et al. Unifying the epidemiological and evolutionary dynamics of pathogens. **Science (New York, N.Y.)**, New York, v. 303. N. 5656, p. 327- 32, jan. 2004. doi: 10.1126/science.1090727.

GU, J.; KORTEWEG C. Pathology and pathogenesis of severe acute respiratory syndrome. **The American journal of pathology**, Philadelphia, v. 170, n. 4, p. 1136-47, abr. 2007. doi: 10.2353/ajpath.2007.061088.

HADFIELD, J. et al. Nextstrain: real-time tracking of pathogen evolution. **Bioinformatics**, Oxford, v. 34, n. 23, p. 4121-4123, dez. 2018. doi: 10.1093/bioinformatics/bty407.

HALACHEV, M. et al. Increased ultra-rare variant load in an isolated Scottish population impacts exonic and regulatory regions. **PLoS genetics**, San Francisco, v. 15, n. 11, p. e1008480, nov. 2019. doi: 10.1371/journal.pgen.1008480.

HAN, Z. et al. Genetic, antigenic and pathogenic characterization of avian coronaviruses isolated from pheasants (*Phasianus colchicus*) in China. **Veterinary microbiology**, Amsterdam, v. 240, jan. 2020. doi: 10.1016/j.vetmic.2019.108513.

HASSAN, S. S. et al. Molecular conservation and differential mutation on ORF3a gene in Indian SARS-CoV2 genomes. **Genomics**, San Diego, v. 112, n. 5, p. 3226-3237, set. 2020. doi: 10.1016/j.ygeno.2020.06.016.

HE, R. et al. The clinical course and its correlated immune status in COVID-19 pneumonia. **Journal of clinical virology: the official publication of the Pan American society for clinical virology**, Amsterdam, v. 127. p. 104361, abr. 2020. doi: 10.1016/j.jcv.2020.104361.

HOFFMANN, M. et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. **Cell**, Cambridge, v. 181, n. 2, p. 271–280.e8, abr. 2020. doi: 10.1016/j.cell.2020.02.052.

HU, B. et al. Characteristics of SARS-CoV-2 and COVID-19. **Nature reviews. Microbiology**, London, v. 6, p. 1-14, out. 2020. doi: 10.1038/s41579-020-00459-7.

HU, B. et al. 2017. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. **PLoS pathogens**, San Francisco, v. 13, n. 11, p. e1006698 nov. 2017. doi: 10.1371/journal.ppat.1006698.

HUSSAIN, S. et al. Identification of Novel Subgenomic RNAs and Noncanonical Transcription Initiation Signals of Severe Acute Respiratory Syndrome Coronavirus. **Journal of virology**, Baltimore, v. 79, n. 9, p. 5288–5295, mai. 2005. doi: 10.1128/JVI.79.9.5288-5295.2005.

HUSSAIN, S. et al. Identification of Novel Subgenomic RNAs and Noncanonical Transcription Initiation Signals of Severe Acute Respiratory Syndrome Coronavirus. **Journal of virology**, Baltimore, v. 79, n. 9, p. 5288–5295, mai. 2005. doi: 10.1128/JVI.79.9.5288-5295.2005.

JAIMES, J. A. et al. Phylogenetic Analysis and Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and Proteolytically Sensitive Activation Loop. **Journal of Molecular Biology**, London, v. 432, n. 10, p. 3309–3325, mai. 2020. doi: 10.1016/j.jmb.2020.04.009.

Ji, W. et al. Cross-species transmission of the newly identified coronavirus 2019-nCoV. **Journal of medical virology**, New York, v. 92, n. 4, p. 433–440, abr. 2020. doi: 10.1002/jmv.25682.

JONES, L. R.; MANRIQUE, J. M. Quantitative phylogenomic evidence reveals a spatially structured SARS-CoV-2 diversity. **Virology**, New York, v. 550, p. 70-77, ago. 2020. doi: 10.1016/j.virol.2020.08.010.

KADAM, S. B. et al. SARS-CoV-2, the pandemic coronavirus: Molecular and structural insights. **Journal of basic microbiology**, Berlin, v. 61, n. 3, p. 180-202, jan. 2021. doi: 10.1002/jobm.202000537.

KARCZEWSKI, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. **Nature**, London, v. 581, n. 7809, p. 434-443, mai. 2020. doi: 10.1038/s41586-020-2308-7.

KAWAGUCHI, S. et al. Comprehensive HLA Typing from a Current Allele Database Using Next-Generation Sequencing Data. **Methods in molecular biology (Clifton, N.J.)**, Clifton, v. 1802, p. 225-233, jun. 2018. doi: 10.1007/978-1-4939-8546-3\_16.

KAWAGUCHI, S.; HIGASA, K. et al. Comprehensive HLA Typing from a Current Allele Database Using Next-Generation Sequencing Data. **Methods in molecular biology (Clifton, N.J.)**, Clifton, v. 1802, p. 225-233, jun. 2018. doi: 10.1007/978-1-4939-8546-3\_16.

KEHDY, F. S. G. et al. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. **Proceedings of the national academy of sciences of the United States of America**, Washington, v. 112, n. 28, p. 8696-701, jul. 2015. doi: 10.1073/pnas.1504447112.

KINDLER, E.; THIEL V. SARS-CoV and IFN: Too Little, Too Late. **Cell host & microbe**, Cambridge, v. 19, n. 2, p. 139-141, fev. 2016. doi: 10.1016/j.chom.2016.01.012.

KOYAMA, T.; PLATT, D.; PARIDA, L. Variant analysis of SARS-CoV-2 genomes. **Bulletin of the World Health Organization**, Geneva, v. 98, n. 7, p. 495–504, jul. 2020. doi: 10.2471/BLT.20.253591.

KUMAR, M.; KHODOR, S. A. Pathophysiology and treatment strategies for COVID-19. **Journal of translational medicine**, v. 18, n. 353, set. 2020. <https://doi.org/10.1186/s12967-020-02520-8>.

LATIF, A. A; MUKARATIRWA, S. Zoonotic origins and animal hosts of coronaviruses causing human disease pandemics: A review. **The onderstepoort journal of veterinary research**, Union Of South Africa, v. 87, n. 1, p. 1895, dez. 2020. doi: 10.4102/ojvr.v87i1.1895.

LAURING, A. S.; ANDINO, R. Quasispecies theory and the behavior of RNA viroses. **PLoS pathogens**, San Francisco, v. 6, n. 7, p. e1001005, jul. 2020. doi: 10.1371/journal.ppat.1001005.

LI, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. **Annual Review of Virology**, Palo Alto, v. 3, n. 1, p. 237-261, ago. 2016. doi: 10.1146/annurev-virology-110615-042301.

LI, X. et al. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. **Journal of medical virology**, New York, v. 10, p. 10.1002/jmv.25731, mar. 2020. doi: 10.1002/jmv.25731.

LI, D. et al. Coronavirus infections and immune responses. **Journal of medical virology**, New York, v. 92, n. 4, p. 424-432, fev. 2020. doi: 10.1002/jmv.25685.

LIMA-COSTA, M. F. et al. Genomic ancestry and ethnoracial self-classification based on 5,871 community-dwelling Brazilians (The Epigen Initiative). **Scientific reports**, London, v.5, p. 9812, abr. 2015. doi: 10.1038/srep09812.

LIU, Y. et al. Correlation Between Relative Nasopharyngeal Virus RNA Load and Lymphocyte Count Disease Severity in Patients with COVID-19. **Viral immunology**, New York, abr. 2020. doi: 10.1089/vim.2020.0062.

LIU, Z. et al. Lymphocyte subset (CD4+, CD8+) counts reflect the severity of infection and predict the clinical outcomes in patients with COVID-19. **The journal of infection**, London, v. 81, n. 2, p. 318-356, abr. 2020. doi: 10.1016/j.jinf.2020.03.054.

LUDWIG, S; ZARBOCK, A. Coronaviruses and SARS-CoV-2: A Brief Overview. **Anesthesia and Analgesia**, Cleveland, v. 131, n. 1, p. 93-96, jul. 2020. doi: 10.1213/ANE.0000000000004845.

LUIS, A. D. et al. A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special?. **Proceedings. biological sciences**, London, v. 280, n. 1756, p. 20122753, fev. 2013. doi: 10.1098/rspb.2012.2753.

LUK, H. K. H. et al. Molecular epidemiology, evolution and phylogeny of SARS coronavirus. **Infection, genetics and evolution: Journal of molecular epidemiology and evolutionary genetics in infectious diseases**, Amsterdam, v. 71, p. 21-30, jul. 2019. doi: 10.1016/j.meegid.2019.03.001.

MAGRONE, T.; MAGRONE, M.; JIRILLO, E. Focus on Receptors for Coronaviruses with Special Reference to Angiotensin- Converting Enzyme 2 as a Potential Drug Target - A Perspective. **Endocrine, metabolic & immune disorders drug targets**, Saif Zone, v. 2020, n. 6, p. 807-811, jul. 2020. doi: 10.2174/1871530320666200427112902.

MANFREDONIA, I. et al. Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. **Nucleic acids research**, London, v. 48, n. 22, p. 12436-12452, dez. 2020. doi: 10.1093/nar/gkaa1053.

MASTERS, P. S. The Molecular Biology of Coronaviruses. **Advances in virus research**, New York, v. 66, p. 193-292, jul. 2006. doi: 10.1016/S0065-3527(06)66005-3.

MENACHERY, V. D. et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. **Nature medicine**, New York, v. 21, n. 12, p. 1508-13, dez. 2015. doi: 10.1038/nm.3985.

MORSE, J. S. et al. Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV. **Chembiochem: a European journal of chemical biology**, Weinheim, v. 21, n. 5, p. 730–738, mar. 2020. doi: 10.1002/cbic.202000047.

MOURA, R. R. et al. Meta-analysis of Brazilian genetic admixture and comparison with other Latin America countries. **American journal of human biology**, New York, v. 27, n. 5, p. 674-680, mar. 2015. Disponível em: <https://doi.org/10.1002/ajhb.22714> . Acesso em: 10 mai 2021.

MOUSAVIZADEH, L.; GHASEMI, S. Genotype and phenotype of COVID-19: Their roles in pathogenesis. **Journal of microbiology, immunology, and infection**, Taipei, mar. 2020. doi: 10.1016/j.jmii.2020.03.022.

NAQVI, A. A. T. et al. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. **Biochimica et biophysica acta. Molecular basis of disease**, Amsterdam, v. 1866, n. 10, p. 165878, out. 2020. doi: 10.1016/j.bbadis.2020.165878.

NASAB, M. G.; SAGHAZADEH A.; REZAEI N. SARS-CoV-2—A Tough Opponent for the Immune System. **Archives of medical research**, Mexico, v. 51, n. 6, p. 589-592, mai. 2020. doi: 10.1016/j.arcmed.2020.05.020.

NASLAVSKY, M. S. et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database. **Human mutation**, New York, v. 38, n. 7, p. 751-763, jul. 2017. doi: 10.1002/humu.23220.

NELEMANS, T.; KIKKERT M. Viral Innate Immune Evasion and the Pathogenesis of Emerging RNA Virus Infections. **Viruses**, Basel, v. 11, n. 10, p. 961, out. 2019. doi: 10.3390/v111100961.

OUASSOU, H. et al. The Pathogenesis of Coronavirus Disease 2019 (COVID-19): Evaluation and Prevention. **Journal of immunology research**, Cairo, v. 2020, jul. 2020. doi: 10.1155/2020/1357983.

PLOWRIGHT, R. K. Pathways to zoonotic spillover. **Nature reviews. Microbiology**, London, v. 15, n. 8, p. 502-510, mai. 2017. doi: 10.1038/nrmicro.2017.45.

PRATHIVIRAJ, R.; KIRAN G. S.; SELVIN J. Phylogenomic proximity and comparative proteomic analysis of SARS-CoV-2. **Gene reports**, New York, v. 20, p. 100777, set. 2020. doi: 10.1016/j.genrep.2020.100777.

QAMAR, M. et al. Structural basis of SARS-CoV-2 3CL<sup>pro</sup> and anti-COVID-19 drug discovery from medicinal plants. **Journal of Pharmaceutical Analysis**, Xi'an, v. 10, n. 4, p. 313–319, ago. 2020. doi: 10.1016/j.jpha.2020.03.009.

QIN, C. et al. Dysregulation of immune response in patients with COVID-19 in Wuhan, China. **Clinical infectious diseases: an official publication of the Infectious Diseases Society of America**, Chicago, v. 71, n. 15, p. 762-768, jul. 2020. doi: 10.1093/cid/ciaa248.

RAGHAV, S. et al. Analysis of Indian SARS-CoV-2 Genomes Reveals Prevalence of D614G Mutation in Spike Protein Predicting an Increase in Interaction With TMPRSS2 and Virus Infectivity. **Frontiers in microbiology**, Lausanne, v. 11, p. 594928, nov. 2020. doi: 10.3389/fmicb.2020.594928.

RAMBAULT, A. et al. The genomic and epidemiological dynamics of human influenza A virus. **Nature**, London, v. 453, n. 7195, p. 615-619, abr. 2008. doi: 10.1038/nature06945.

RISSANEN, I. et al. Idiosyncratic Mòjiāng virus attachment glycoprotein directs a host-cell entry pathway distinct from genetically related henipaviruses. **Nature Communications**, London, v. 8, p. 16060, jul. 2017. doi: 10.1038/ncomms16060.

ROCHA, C. S. et al. The Brazilian Initiative on Precision Medicine (BIPMed): fostering genomic data-sharing of underrepresented populations. **NPJ genomic medicine**, London, v. 5, p. 42, out. 2020. doi: 10.1038/s41525-020-00149-6.

ROCHA, C. S.; SECOLIN, R. et al. The Brazilian Initiative on Precision Medicine (BIPMed): fostering genomic data-sharing of underrepresented populations. **NPJ genomic medicine**, London, v. 5, p. 42, out. 2020. doi: 10.1038/s41525-020-00149-6.

ROGNAN, D. et al. Molecular dynamics study of a complex between the human histocompatibility antigen HLA-A2 and the IMP58-66 nonapeptide from influenza virus matrix protein. **European journal of biochemistry**, Berlin, v. 208, n. 1, p. 101-13, ago. 1992. doi: 10.1111/j.1432-1033.1992.tb17163.x.

SAHA, I. et al. Genome-wide analysis of Indian SARS-CoV-2 genomes for the identification of genetic mutation and SNP. **Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases**, Amsterdam, v. 85, p. 104457, jul. 2020. doi: 10.1016/j.meegid.2020.104457.

SALLARD, E. et al. Tracing the origins of SARS-COV-2 in coronavirus phylogenies: a review. **Environmental chemistry letters**, Secaucus, v. 4, p. 1–17, fev. 2021. doi: 10.1007/s10311-020-01151-1.

SAWICKI, S. G. et al. A Contemporary View of Coronavirus Transcription. **Journal of virology**, Baltimore, v. 81, n. 1, p. 20–29, jan. 2007. doi: 10.1128/JVI.01358-06.

SCHOLZ, J. R. COVID-19, Sistema Renina-Angiotensina, Enzima Conversora da Angiotensina 2 e Nicotina: Qual a Inter-Relação?. **Arquivos Brasileiros de Cardiologia**, Rio de Janeiro, v. 115, n. 4, out. 2020. <https://doi.org/10.36660/abc.20200653>.

SCHWARZ, J. M. et al. MutationTaster evaluates disease-causing potential of sequence alterations. **Nature methods**, New York, v. 7, n. 8, p. 575-6, ag. 2010. doi: 10.1038/nmeth0810-575.

SEAH, I.; AGRAWAL, R. Can the Coronavirus Disease 2019 (COVID-19) Affect the Eyes? A Review of Coronaviruses and Ocular Implications in Humans and Animals. **Ocular immunology and inflammation**, Buren, v. 28, p. 1-5, mar. 2020. doi: 10.1080/09273948.2020.1738501.

SECOLIN, R. et al. Genetic variability in COVID-19-related genes in the Brazilian population. **Nature**, London, v. 8, abr. 2021. Disponível em: <https://doi.org/10.1038/s41439-021-00146-w>. Acesso em: 14 mai 2021.

SECOLIN, R.; MAS-SANDOVAL, A. et al. Distribution of local ancestry and evidence of adaptation in admixed populations. **Scientific reports**, London, v. 9, n. 1, p. 13900, set. 2019. doi: 10.1038/s41598-019-50362-2.

SHEN, Z. et al. A conserved region of nonstructural protein 1 from alphacoronaviruses inhibits host gene expression and is critical for viral virulence. **The journal of biological chemistry**, Baltimore, v. 294, n. 37, p. 13606–13618, set. 2019. doi: 10.1074/jbc.RA119.009713.

SHEREEN, M. A. et al. COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. **Journal of advanced research**, Giza, v. 24, p. 91–98, jul. 2020. doi: 10.1016/j.jare.2020.03.005.

SMITH, E. C. et al. Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. **PLoS pathogens**, v. 9, n. 8, p. e1003565, ago. 2013. doi: 10.1371/journal.ppat.1003565.

SNIJDER, E. J. et al. Ultrastructure and Origin of Membrane Vesicles Associated with the Severe Acute Respiratory Syndrome Coronavirus Replication Complex. **Journal of virology**, Baltimore, v. 80, n. 12, p. 5927–5940, jun. 2006. doi: 10.1128/JVI.02501-05.

SONG, Z. et al. From SARS to MERS, Thrusting Coronaviruses into the Spotlight. **Viruses**, Basel, v. 11, n. 1, p. 59, jan. 2019. doi: 10.3390/v11010059.

SRINIVASAN, S. et al. Structural Genomics of SARS-CoV-2 Indicates Evolutionary Conserved Functional Regions of Viral Proteins. **Viruses**, Basel, v. 12, n. 4, p. 360, abr. 2020. doi: 10.3390/v12040360.

STERN A.; ANDINO R. Viral evolution: it is all about mutations. **Viral pathogenesis**, New York, p. 233-40, fev. 2016. Disponível em: <https://doi.org/10.1016/B978-0-12-800964-2.00017-3>. Acesso em: 2 mai 2021.



SU, S. et al. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. **Trends in microbiology**, Cambridge, v. 24. N. 6, p. 490-502, mar. 2016. doi: 10.1016/j.tim.2016.03.003.

SUBUDHI, S.; RAPIN, N.; MISRA, V. Immune System Modulation and Viral Persistence in Bats: Understanding Viral Spillover. **Viruses**, Basel, v. 11, n. 2, p. 192, fev. 2019. doi: 10.3390/v11020192.

SUTTON, D. et al. Universal Screening for SARS-CoV-2 in Women Admitted for Delivery. **The New England journal of medicine**, Boston, v. 382, n.22, p. 2163-2164, mai. 2020. doi: 10.1056/NEJMc2009316.

TANG, X. et al. On the origin and continuing evolution of SARS-CoV-2. **National science review**, Beijing, p. nwaa036, mar. 2020. doi: 10.1093/nsr/nwaa036.

THEVARAJAN, I.; BUISING, K. L.; COWIE, B. C. Clinical presentation and management of COVID-19. **The medical journal of Australia**, Sydney, v. 213, n. 3, p. 134–139, ago. 2020. doi: 10.5694/mja2.50698.

UDDIN, M. et al. SARS-CoV-2/COVID-19: Viral Genomics, Epidemiology, Vaccines, and Therapeutic Interventions. **Viruses**, Basel, v. 12, n. 5, p. 526, mai. 2020. doi: 10.3390/v12050526.

WANG, F. et al. Characteristics of Peripheral Lymphocyte Subset Alteration in COVID-19 Pneumonia. **The journal of infectious diseases**, Chicago, p. jiaa150, mar. 2020. doi: 10.1093/infdis/jiaa150.

WANG, D. et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. **JAMA**, Chicago, v. 323, n. 11, p. 1061–1069, mar. 2020. doi: 10.1001/jama.2020.1585.

WEBER, S.; RAMIREZ, C.; DOERFLER, W. Signal hotspot mutations in SARS-CoV-2 genomes evolve as the virus spreads and actively replicates in different parts of the world. **Virus research**, Amsterdam, v. 289, p. 198170, set. 2020. doi: 10.1016/j.virusres.2020.198170.

WONG, L. Y. R; LUI, P. Y.; JIN, D.Y. A molecular arms race between host innate antiviral response and emerging human coronaviruses. **Virologica sinica**, Beijing, v. 31, n. 1, p. 12–23, fev 2016. doi: 10.1007/s12250-015-3683-3.

WU, A. et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. **Cell host & microbe**, Cambridge, v. 27, n. 3, p. 325–328, fev. 2020. doi: 10.1016/j.chom.2020.02.001.

WU, D. et al. The SARS-CoV-2 outbreak: What we know. **International society for infectious diseases**, Hamilton, v. 94, p. 44–48, mar. 2020. doi: 10.1016/j.ijid.2020.03.004.

WU, Z. et al. Novel Henipa-like virus, Mojiang Paramyxovirus, in rats, China, 2012. **Emerging infectious diseases**, Atlanta, v. 20, n. 6, p. 1064-6, jun. 2014. doi: 10.3201/eid2006.131022.

XIA, S. et al. Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. **Cell research**, Beijing, v. 30, n. 4, p. 343–355, abr. 2020. doi: 10.1038/s41422-020-0305-x.

XIONG, Y. et al. Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. **Emerging microbes & infections**, New York, v. 9, n. 1, p. 761-770, mar. 2020. doi: 10.1080/22221751.2020.1747363.

YADAV, P. D. et al. Full-genome sequences of the first two SARS-CoV-2 viruses from India. **The Indian journal of medical research**, New Delhi, v. 151, n. 2-3, p. 200-209, fev-mar. 2020. doi: 10.4103/ijmr.IJMR\_663\_20.

YE, Z. W. et al. Zoonotic origins of human coronaviruses. **International journal of biological sciences**, Lake Haven, v. 16, n. 10, p. 1686–1697, mar. 2020. doi: 10.7150/ijbs.45472.

ZAKI, A. M. et al. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. **The New England journal of medicine**, Boston, v. 367, n. 19, p. 1814-20, nov. 2012. doi: 10.1056/NEJMoa1211721.

ZHANG, Q. et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. **Science**, v. 370, n. 6515, p. eabd4570, out. 2020. doi: 10.1126/science.abd4570.

ZHANG, T.; WU Q.; ZHANG Z. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. **Current biology: CB**, London, v. 30, n. 7, p. 1346-1351.e2, mar. 2020. doi: 10.1016/j.cub.2020.03.022.

ZHOU, H. et al. A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein. **Current biology: CB**, London, v. 30, n. 11, p. 2196-2203.e3, jun. 2020. doi: 10.1016/j.cub.2020.05.023.