



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"

**FELIPE ALLAN DA SILVA DA COSTA**

**Análise da dispersão temporal e epidemiológica das variantes do SARS-CoV-2  
no período de pré-vacinação em massa na cidade de Botucatu-SP**

**Araraquara**

**2022**

**FELIPE ALLAN DA SILVA DA COSTA**

**Análise da dispersão temporal e epidemiológica das variantes do SARS-CoV-2  
no período de pré-vacinação em massa na cidade de Botucatu-SP**

Dissertação apresentada ao Programa de Pós-graduação em Biociências e Biotecnologia Aplicadas à Farmácia, da Faculdade de Ciências Farmacêuticas, UNESP, como parte dos requisitos para obtenção do Título de Mestre, área de pesquisa em Biologia Molecular.

Orientador: Prof<sup>o</sup>. Dr<sup>o</sup>. Jayme Augusto Souza Neto

Coorientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Rejane Maria Tommasini Grotto

Coorientador: Prof<sup>o</sup>. Dr<sup>o</sup>. Carlos Magno Castelo Branco

Fortaleza

**Araraquara**

**2022**

---

**C837a** Costa, Felipe Allan da Silva da.  
Análise da dispersão temporal e epidemiológica das variantes do SARS-CoV-2 no período de pré-vacinação em massa na cidade de Botucatu-SP / Felipe Allan da Silva da Costa. – Araraquara: [S.n.], 2022.  
63 f. : il.

Dissertação (Mestrado) – Universidade Estadual Paulista. “Júlio de Mesquita Filho”. Faculdade de Ciências Farmacêuticas. Programa de Pós Graduação em Biociências e Biotecnologia Aplicadas à Farmácia. Área de Biologia Molecular.

Orientador: Jayme Augusto de Souza Neto.  
Coorientadora: Rejane Maria Tommasini Grotto.  
Coorientador: Carlos Magno Castelo Branco Fortaleza.

1. SARS-CoV-2. 2. Variantes do SARS-COV-2. 3. Epidemiologia Molecular. I. Souza-Neto, Jayme Augusto de, orient. II. Grotto, Rejane Maria Tommasini, coorient. III. Fortaleza, Carlos Magno Castelo Branco, coorient. IV. Título.

## **CERTIFICADO DE APROVAÇÃO**

**TÍTULO DA DISSERTAÇÃO:** Análise da dispersão temporal e epidemiológica das variantes do SARS-CoV-2 no período de pré-vacinação em massa na cidade de Botucatu-SP

**AUTOR:** FELIPE ALLAN DA SILVA DA COSTA

**ORIENTADOR:** JAYME AUGUSTO DE SOUZA NETO

**COORIENTADORA:** REJANE MARIA TOMMASINI GROTTTO

**COORIENTADOR:** CARLOS MAGNO CASTELO BRANCO FORTALEZA

Aprovado como parte das exigências para obtenção do Título de Mestre em Biociências e Biotecnologia Aplicadas à Farmácia, área: Análises Clínicas pela Comissão Examinadora:

Prof. Dr. JAYME AUGUSTO DE SOUZA NETO (Participação Virtual)  
Departamento de Bioprocessos e Biotecnologia / Faculdade de Ciências Agronômicas de Botucatu - Unesp

Profa. Dra. MARIA CAROLINA QUARTIM BARBOSA ELIAS-SABBAGA (Participação Virtual)  
Departamento de Parasitologia / Instituto Butantã de São Paulo

Dra. KAREN INGRID TASCA (Participação Virtual)  
Faculdade de Medicina de Botucatu - Unesp

Araraquara, 27 de setembro de 2022

A minha mãe Luciene, a minha Tia Elaine e a minha irmã Maria Eduarda por todo amor e compreensão.

## **AGRADECIMENTOS**

Agradeço primeiramente ao apoio financeiro e emocional da minha família, especialmente minha mãe Luciene, minha tia Elaine e minha irmã Maria Eduarda, mulheres inspiradoras que não mediram esforços para que eu chegasse e concluísse essa etapa da minha vida.

Ao meu namorado Tiago Rodrigo Beloti por todo apoio, amor e carinho durante essa jornada.

Agradeço ao povo brasileiro por contribuir para o suporte e manutenção das bolsas de estudo, sobretudo àquelas implementadas pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) a qual fui contemplado.

Ao professor e orientador Jayme Augusto Souza Neto pela sua atenção, compreensão e respeito desde o início, além de se mostrar bem acessível e dedicado ao meu aprendizado. Também agradeço pela oportunidade concedida em estagiar no laboratório no Treinamento Técnico, que foi de suma importância para eu descobrir o amor pela área.

A minha coorientadora Professora e Doutora Rejane Maria Tommasini Grotto por estar presente durante todo processo deste trabalho e me incentivar a continuar na área da pesquisa. Ao meu orientador Professor e Doutor Carlos Magno Castelo Branco Fortaleza por auxiliar nas análises epidemiológicas e bioestatísticas, processos fundamentais no andamento dessa pesquisa.

Por último e não menos importante, agradeço as pessoas do Laboratório de Genômica Funcional & Microbiologia de Vetores da Unesp de Botucatu-SP por contribuírem e participarem ativamente dos experimentos e desenvolvimento desta pesquisa. Meu muito obrigado, Patrícia Akemi Assato, Drielle Baptista Dos Santos Figueiredo, Tassia Turini e Leonardo Nazario de Moraes por todo aparato no desenvolvimento deste trabalho e apoio emocion

## RESUMO

Após a circulação das Variantes de Preocupação (VOC) do SARS-CoV-2, Alpha, Beta, Gamma, Delta e a atual predominante Ômicron, e a ampla cobertura vacinal, houve uma drástica redução nos números de óbitos por Covid-19. Em maio de 2021, o município de Botucatu-SP foi palco de uma pesquisa que avaliou a efetividade da Vacina COVID-19 Recombinante/Fiocruz® contra variantes do SARS-CoV-2, possibilitando a imunização de mais de 60 mil pessoas em um único dia. **Objetivos:** O objetivo central do presente estudo foi avaliar a dispersão temporal e epidemiológica das variantes do SARS-CoV-2 antes da vacinação massiva que ocorreu em Botucatu, além de relacioná-las às características clínicas da doença. **Metodologia:** Foram selecionadas aleatoriamente 400 amostras positivas para SARS-CoV-2, referentes as 4 Semanas Epidemiológicas (SE 16, 17, 18 e 19 de 2021) que antecederam o dia da vacinação. O Sequenciamento de Nova Geração foi utilizado para produzir as sequências e determinar as variantes. **Resultados:** Foram geradas 371 sequências de alta qualidade. Dessas sequências, 98,4% eram da VOC Gamma, enquanto 1,3% eram da VOC Alpha e 0,3% eram da Variante de Interesse Zeta. Durante as 4 SE as linhagens de Gamma P.1 e P.1.14 foram as mais frequentes. A linhagem P.1 foi mais frequente tanto em homens quanto em mulheres e nas pessoas mais jovens e nos adultos (0-59 anos). Não houve correlação entre as variantes e a presença de comorbidades, tampouco entre elas e os quadros clínicos moderados e graves da Covid-19 ou óbito. Entretanto, P.1 foi mais frequente do que P.1.14 em pessoas que apresentaram algum sintoma da Covid-19. Não houve diferença entre os valores de PCR em tempo real ( $C_T$ ) para essas linhagens. Nas análises filogenéticas, foi possível observar um pequeno agrupamento de 9 amostras de P.1.14 contendo as mutações em ORF1a: M584V e A3620V, até então, não encontradas em nenhuma sequência dessa sublinhagem. **Conclusão:** A alta predominância de P.1 e P.1.14 em um cenário pré-vacinação em massa pode nos fornecer *insights* sobre a evolução e a epidemiologia molecular do SARS-CoV-2 e suas VOCs. Dessa forma ressaltamos a importância da vigilância genômica do SARS-CoV-2, que pode ajudar a subsidiar as tomadas de decisões dos setores públicos e manejo da Covid-19.

**Palavras-chave:** SARS-CoV-2; Variantes do SARS-CoV-2; Epidemiologia Molecular.

## ABSTRACT

After the circulation of the Variants of Concern (VOC) of SARS-CoV-2, Alpha, Beta, Gamma, Delta and the current predominant Ômicron, and the wide vaccine coverage, there was a drastic reduction in the numbers of deaths by Covid-19. In May 2021, the municipality of Botucatu-SP was the scene of a survey that evaluated the effectiveness of the Vaccine COVID-19 Recombinant/Fiocruz® against variants of SARS-CoV-2, enabling the immunization of more than 60,000 people in a single day. Objectives: The main objective of the present study was to evaluate the temporal and epidemiological dispersion of the SARS-CoV-2 variants before the massive vaccination that took place in Botucatu, in addition to relating them to the clinical characteristics of the disease. Methodology: 400 samples positive for SARS-CoV-2 were randomly selected, referring to the 4 Epidemiological Weeks (SE 16, 17, 18 and 19 of 2021) that preceded the day of vaccination. Next Generation Sequencing was used to generate the sequences and determine the variants. Results: 371 high quality sequences were generated. Of these sequences, 98.4% were from VOC Gamma, while 1.3% were from VOC Alpha and 0.3% were from Variant of Interest Zeta. During the 4 SEs, the Gamma P.1 and P.1.14 strains were the most frequent. Lineage P.1 was more frequent in both men and women and in younger people and adults (0-59 years). There was no correlation between the variants and the presence of comorbidities, nor between them and the moderate and severe clinical conditions of Covid-19 or death. However, P.1 was more frequent than P.1.14 in people who had some symptoms of Covid-19. There was no difference between real-time PCR (CT) values for these strains. In the phylogenetic analyses, it was possible to observe a small grouping of 9 P.1.14 samples containing the mutations in ORF1a: M584V and A3620V, until then, not found in any sequence of this subline. Conclusion: The high prevalence of P.1 and P.1.14 in a pre-mass vaccination scenario may provide us with insights into the evolution and molecular epidemiology of SARS-CoV-2 and its VOCs. In this way, we emphasize the importance of SARS-CoV-2 genomic surveillance, which can help support public sector decision-making and the management of Covid-19.

**Keywords:** SARS-CoV-2; SARS-CoV-2 variants; Molecular Epidemiology.



## LISTA DE ILUSTRAÇÕES

|                                                                                                                                                                                                                                                                              |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1: Distribuição temporal das variantes do SARS-CoV-2 nas SE 16, 17, 18 e 19 de 2021 no município de Botucatu-SP. ....                                                                                                                                                 | 28 |
| Figura 2: Distribuição das sublinhagens da VOC do SARS-CoV-2 Gamma nas SE 16, 17, 18 e 19 de 2021 no município de Botucatu-SP. O valor de p (*) indica a significância estatística encontrada, indicando a maior proporção de P.1 quando comparada com P.1.14 na SE 16 ..... | 30 |
| Figura 3: Proporção das variantes do SARS-CoV-2 encontradas através das SE 16, 17, 18 e 19 de 2021 em relação ao total de casos positivos de cada SE, representados pela linha preta. ....                                                                                   | 31 |
| Figura 4: Linhagens encontradas pela Rede Butantan de Alerta Pandêmico de Variantes SARS-CoV-2 durante a vigilância genômica do SARS-CoV-2 nas SE 16, 17, 18 e 19 de 2021. ....                                                                                              | 32 |
| Figura 5: Departamentos Regionais de Saúde (DRS) do estado de São Paulo. ....                                                                                                                                                                                                | 33 |
| Figura 6: Frequências das linhagens P.1 e P.1.14 nas DRS que circulam o município de Botucatu, incluindo a DRS da capital São Paulo. ....                                                                                                                                    | 34 |
| Figura 7: Distribuição das variantes do SARS-CoV-2 de acordo com cada gênero. O valor de p (*) indica a significância estatística encontrada nas frequências de P.1 quando comparada com a sublinhagem P.1.14. ....                                                          | 36 |
| Figura 8: Frequência das variantes do SARS-CoV-2 distribuídas em diferentes grupos de acordo com as faixas etárias. O "n" indica o número total de amostras sequenciadas em cada grupo. O valor de p (*) indica a significância estatística. ....                            | 37 |
| Figura 9: Frequências das variantes do SARS-CoV-2 entre as pessoas com diferentes tipos de comorbidades. O "n" indica o número de participantes deste estudo que convivem com determinada comorbidade. ....                                                                  | 38 |
| Figura 10: Árvore filogenética com as amostras de Botucatu-SP coletadas antes da vacinação em massa destacadas em amarelo. ....                                                                                                                                              | 40 |
| Figura 11: Árvore filogenética destacando o clado 20B contendo todas as 371 amostras do presente estudo realizado em Botucatu-SP nas SE 16, 17, 18 e 19 de 2021. ....                                                                                                        | 41 |
| Figura 12: Árvore filogenética contendo apenas as linhagens de Gamma. Dentro do triângulo vermelho em destaque, estão as sequências da sublinhagem P.1.14 agrupadas no clado contendo a substituição em ORF1a:S2947N. As outras amostras são majoritariamente P.1. ....      | 42 |

Figura 13: Ramo da árvore da sublinhagem P.1.14 que contém a mutação em ORF1a:S2947N. ....43

Figura 14: Trecho da árvore filogenética destacando o ramo com amostras deste estudo que se agruparam no ramo contendo as substituições em ORF1a: M584V e A3620V.44

## LISTA DE TABELAS

|                                                                                                                                                                                                                         |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela 1: Relação de amostras positivas e sequenciadas nas SE selecionadas do estudo. Encontram-se também a quantidade de genomas que passaram em todos os testes de qualidade e as plataformas de sequenciamento. .... | 25 |
| Tabela 2: Relação das comorbidades importantes para Covid-19 encontradas neste estudo e a quantidade de participantes que convivem com essas doenças.....                                                               | 27 |
| Tabela 3: Proporção da distribuição das variantes do SARS-CoV-2 do total de sequências produzidas em cada gênero. ....                                                                                                  | 35 |

## LISTA DE ABREVIATURAS E SIGLAS

**cDNA** – DNA Complementar

**CEP** – Comitê de Ética em Pesquisa

**Covid-19** - Doença Pelo Coronavírus de 2019

**DRS** – Departamentos Regionais de Saúde

**ECA-2** – Enzima Conversora da Angiotensina 2

**FMB** – Faculdade de Medicina de Botucatu

**GISAID** - Iniciativa Global de Compartilhamento de Dados de Influenza Aviária

**HC** – Hospital das Clínicas

**HCFMB** - Hospital das Clínicas da Faculdade de Medicina de Botucatu

**MERS-CoV** - Coronavírus da Síndrome Respiratória do Oriente Médio

**nsp** – Proteína não Estrutural

**OMS** – Organização Mundial da Saúde

**ORF** - Quadro de Leitura Aberto

**PCR** - Reação em Cadeia da Polimerase

**pM** – Picomolar

**qPCR** - Reação em Cadeia da Polimerase Quantitativa

**RBD** – Domínio de Ligação ao Receptor

**RdRp** - RNA polimerase dependente de RNA

**RNA** – Ácido Ribonucleico

**RT-qPCR** - Transcrição Reversa seguida de Reação em Cadeia da Polimerase Quantitativa

**SARS-CoV** - Coronavírus da Síndrome Respiratória Aguda Grave

**SARS-CoV-2** - Coronavírus da Síndrome Respiratória Aguda Grave 2

**SE** – Semana Epidemiológica

**UNESP** - Universidade Estadual Paulista

**UTI** – Unidade de Terapia Intensiva

**VOC** – Variante de Preocupação

**VOI** – Variante de Interesse

## SUMÁRIO

|                                                                                                                     |           |
|---------------------------------------------------------------------------------------------------------------------|-----------|
| <b>1 INTRODUÇÃO.....</b>                                                                                            | <b>14</b> |
| 1.1 Objetivos.....                                                                                                  | 18        |
| 1.1.1 Geral.....                                                                                                    | 18        |
| 1.1.2 Específicos .....                                                                                             | 18        |
| <b>2 METODOLOGIA.....</b>                                                                                           | <b>20</b> |
| 2.1 Seleção das amostras para o estudo .....                                                                        | 20        |
| 2.2 Obtenção das amostras positivas para SARS-CoV-2 .....                                                           | 20        |
| 2.3 Extração de RNA e confirmação da presença do SARS-CoV-2 por RT-qPCR .....                                       | 20        |
| 2.4 Preparo das bibliotecas e Sequenciamento do Genoma do SARS-CoV-2 .....                                          | 21        |
| <b>2.4.1 Preparo das bibliotecas e sequenciamento Illumina.....</b>                                                 | <b>21</b> |
| 2.4.1.1 Fluxo de bioinformática - Illumina.....                                                                     | 22        |
| <b>2.4.2 Preparo das bibliotecas e sequenciamento ION Torrent .....</b>                                             | <b>22</b> |
| 2.4.2.1 Fluxo de bioinformática – <i>Ion Torrent</i> .....                                                          | 23        |
| 2.5 Análise estatísticas dos Resultados.....                                                                        | 23        |
| 2.6 Análises Filogenéticas .....                                                                                    | 24        |
| <b>3 RESULTADOS .....</b>                                                                                           | <b>25</b> |
| 3.1 Amostras sequenciadas e qualidade do sequenciamento.....                                                        | 25        |
| 3.2 Descrição dos dados demográficos, epidemiológicos e clínicos dos participantes do estudo .....                  | 26        |
| 3.3 Distribuição temporal das variantes do SARS-CoV-2 antes da vacinação em massa em Botucatu .....                 | 28        |
| <b>3.3.1 Distribuição das sublinhagens de Gamma do SARS-CoV-2 nas SE 16, 17, 18 e 19 de 2021.....</b>               | <b>29</b> |
| 3.4 Comparação das frequências das linhagens do SARS-CoV-2 da cidade de Botucatu-SP com o estado de São Paulo ..... | 31        |
| 3.5 Distribuição das variantes do SARS-CoV-2 de acordo com os gêneros .....                                         | 34        |
| 3.6 Distribuição das variantes do SARS-COV-2 em diferentes faixas etárias .....                                     | 36        |
| 3.7 Associação das linhagens do SARS-CoV-2 com comorbidades.....                                                    | 37        |
| 3.8 Associação das linhagens do SARS-CoV-2 com a situação clínica e vacinal dos participantes .....                 | 39        |
| 3.9 Análises Filogenéticas .....                                                                                    | 39        |
| <b>4 DISCUSSÃO .....</b>                                                                                            | <b>45</b> |

|                                    |           |
|------------------------------------|-----------|
| <b>5 CONSIDERAÇÕES FINAIS.....</b> | <b>53</b> |
| <b>REFERÊNCIAS .....</b>           | <b>54</b> |

## 1 INTRODUÇÃO

Os coronavírus são vírus envelopados com uma fita simples de RNA de sentido positivo em seu genoma e podem infectar diversos animais, inclusive seres humanos. Estes vírus podem causar infecções respiratórias variando de brandas a graves, como é o exemplo dos Coronavírus da Síndrome Respiratória Aguda Grave (SARS-CoV) e da Síndrome Respiratória do Oriente Médio (MERS-CoV). Estes vírus causaram surtos de casos em 2002-2003 e 2012, respectivamente (ZHONG et al., 2003; ZAKI et al., 2012).

Em dezembro de 2019, foram registrados casos de uma pneumonia de origem desconhecida em Wuhan, na capital da província de Hubei na China, onde a maioria dos pacientes sintomáticos estavam ligados a um mercado atacadista de frutos do mar e animais da cidade (WUHAN MUNICIPAL HEALTH COMMISSION, 2019). Após a investigação epidemiológica e etiológica desses surtos, pesquisadores isolaram e identificaram um novo coronavírus, provisoriamente denominado 2019- nCoV (ZHU et al., 2020), atualmente designado como SARS-CoV-2 (Coronavírus da Síndrome Respiratória Aguda Grave 2) (GORBALENYA et al., 2020). Mesmo após a implementação das medidas de contenção, o vírus se espalhou rapidamente por todo mundo. No dia 11 de março de 2020, a Organização Mundial da Saúde (OMS) declarou a pandemia da Covid-19 (Doença Pelo Coronavírus de 2019), alertando que os números de infectados e mortes deveriam subir em vários países (WHO, 2020a).

Até o presente momento (08/09/2022) foram notificados à OMS 603.711.760 casos positivos e 6.484.136 óbitos causados pela Covid-19 no mundo. No Brasil o cenário também foi profundamente impactado. Até agora foram registrados 34.477.539 casos e 684.425 mortes (WHO, 2022b).

O genoma do SARS-CoV-2 possui aproximadamente 30 mil bases (Lu et al., 2020) incluindo 14 quadros de leitura abertos (ORFs). A maioria das ORFs codificam 16 proteínas não estruturais (nsp 1 a 16) que fazem parte do processo de replicação viral. O restante das ORFs codifica 5 proteínas acessórias (ORF3, ORF4a, ORF4b, ORF5 e ORF8) e também as quatro proteínas estruturais básicas, que incluem as proteínas de envelope (E), de membrana (M), de nucleocapsídeo (N) e a proteína spike (S) (NAQVI et al., 2020; ZHANG et al., 2020).

A proteína S é de suma importância para o processo infeccioso, mediando a ligação do vírus ao receptor celular, a Enzima Conversora da Angiotensina II (ECA2). A proteína S se liga à ECA2 por meio do domínio de ligação ao receptor (RBD), presente na

subunidade S1 da proteína S. Após isso, ocorre a fusão da subunidade S2 à membrana celular (HARRISON; LIN; WANG, 2020; CEVIK et al., 2020).

Assim como outros vírus, o SARS-CoV-2 também sofre diversas mutações em seu genoma. Estudos do genoma do SARS-CoV-2 têm demonstrado a ocorrência de mutações em diferentes isolados em todo mundo. Essas mutações se concentram principalmente no gene codificador da proteína S, sendo o RBD considerado como um *hotspot* mutacional (SHEIK et al., 2020; VAN DORP et al., 2020). Outros locais de variações e *hostpots* mutacionais no genoma do vírus têm sido descritos (WANG et al., 2020), como é o exemplo do gene da RNA polimerase dependente de RNA (RdRp) (PACHIETTI et al., 2020).

Essas mutações que ocorrem no genoma do SARS-CoV-2 bem como em outros vírus, podem resultar em fenótipos distintos aos das primeiras linhagens. Tais fenótipos podem apresentar alterações na epidemiologia do vírus, na sua transmissibilidade, no tropismo celular, na patologia ou no percurso da doença, no escape imunológico ou podem até mesmo dificultar as ferramentas de diagnóstico. Essas mutações, ao longo do tempo, também podem gerar novas linhagens do vírus - também conhecidas como variantes - e são determinantes no percurso evolutivo viral (VIGNUZZI et al., 2006; VIGNUZZI; ANDINO, 2008; COMBE; SANJUÁN, 2014; DUFFY, 2018; LESBON et al., 2021).

A OMS definiu dois grupos de variantes de acordo com a relevância médica para a saúde pública. O primeiro grupo constitui as Variantes de Interesse do SARS-CoV-2 (VOI), as quais devem possuir alterações fenotípicas ou mutações no genoma, em relação ao vírus de referência, que resultam em alterações de aminoácidos. As manifestações fenotípicas devem resultar em impactos na epidemiologia, antigenicidade ou virulência e também sobre os diagnósticos, terapias e vacinas. Além disso, uma VOI também deve ser considerada causadora de múltiplos casos ou ainda, ser avaliada como uma VOI pela OMS e o Grupo de Trabalho de Evolução do Vírus SARS-CoV-2 (WHO, 2021c).

Um outro grupo de variantes de interesses epidemiológicos são as Variantes de Preocupação do SARS-CoV-2 (VOC). Essas variantes, além de serem VOI, também devem estar associadas ao aumento da transmissibilidade ou agravamento na epidemiologia da Covid-19; ao aumento da virulência do patógeno ou alteração do quadro clínico da doença; e aos impactos negativos sobre as medidas sociais adotadas no enfrentamento da pandemia, no diagnóstico, no tratamento da doença e na eficácia das vacinas. Além disso, também pode ser avaliada como uma VOC pela OMS juntamente com o Grupo de Trabalho de Evolução do Vírus SARS-CoV-2 (WHO, 2021c).



As VOC do SARS-CoV-2 até o momento são: Alpha, Beta, Gama, Delta e Omicron. Juntas, estão associadas a maiores taxas de transmissibilidade e ao escape imunológico de vacinas ou de soros convalescentes. Essas questões causam bastante preocupação devido a ameaça da eficácia das vacinas, bem como o surgimento de novas ondas de contágio por causa das variantes que podem escapar da neutralização desencadeada por infecções anteriores. Sobre a relação entre variantes e diferentes aspectos clínicos da Covid-19 ainda não se tem uma confirmação clara (RAMAN; PATEL; RANJAN, 2021; LAZAREVIC et al., 2021; ALEEM; SAMAD; SLENKER, 2021; BOEHM et al., 2021; PAPANIKOLAOU et al., 2022).

Em relação à eficácia das vacinas contra as VOCs do SARS-CoV-2, estudos têm demonstrado bons resultados dos imunizantes disponíveis, entretanto alguns pontos são importantes de serem levantados. A VOC Alpha mostrou reduções significativas na neutralização por soros de vacinados pela BNT162b2 (BioNTech-Pfizer), variando de 1,5x a 6,7x (BATES et al., 2021; COLLIER et al., 2021; GARCIA-BELTRAN et al., 2021; KUZMINA et al., 2021). Quanto a eficácia clínica, os resultados foram muito positivos para a vacina ChAdOx1-nCov-19 - AZD1222 (AstraZeneca). Esse imunizante demonstrou uma eficácia contra a VOC Alpha de 81% para casos sintomáticos e 73% contra hospitalizações (NASREEN et al., 2021).

A variante Beta, em especial, causou grande preocupação acerca do escape imunológico. Estudos têm demonstrado reduções nos títulos de neutralização de soros vacinados variando de 4,2x até 41,2x (ANICHINI et al., 2021; DEJNIRATTISAI et al., 2021; GARCIA BELTRAN et al., 2021; WALL et al., 2021). A eficácia clínica contra essa VOC também foi significativamente menor (< 50%) para a vacina NVX-CoV2373 (Novavax) (CALLAWAY; MALLAPATY, 2021; NOVAVAX, 2021a).

Para Gamma, a redução dos títulos de neutralização em soros vacinados pela BNT162b2 (Pfizer) varia de 2,3x a 6,7x (GARCIA-BELTRAN et al., 2021; LUSTIG et al., 2021). E a eficácia clínica pelo mesmo imunizante, chegou a 89% contra casos sintomáticos e 95% contra hospitalizações e óbitos (NASREEN et al., 2021). Os títulos de neutralização contra a VOC Delta a partir de soros vacinados pela BNT162b2 (BioNTech-Pfizer) também foram reduzidos, em 1,54x até 5,8x (PLANAS et al., 2021; WALL et al., 2021). A eficácia clínica dessa vacina contra Delta chegou até 88% contra casos sintomáticos (BERNAL et al., 2021a; NASREEN et al., 2021). Por um outro lado, a eficácia clínica da vacina ChAdOx1-nCov-19 - AZD1222 (AstraZeneca) contra Delta foi menor, chegando a 67% contra casos sintomáticos (BERNAL et al., 2021b).

No caso da mais recente VOC, a Omicron, uma alta capacidade de resistência a anticorpos naturais ou induzidos tem sido observada. A redução da neutralização dessa variante, quando comparada com Delta, foi reduzida de 12 a 44 vezes em pacientes convalescentes ou indivíduos vacinados com o imunizante BNT162b2 (BioNTech-Pfizer). Quando se trata de imunização heteróloga, induzida por ChAdOx1-nCov-19 - AZD1222 (AstraZeneca) e BNT162b2 (BioNTech-Pfizer) ou três doses da BNT162b2 (BioNTech-Pfizer), uma maior eficácia é notada. Mesmo nesses cenários, a VOC Omicron ainda assim exibe maior capacidade de fuga imunológica do que Delta (HOFFMANN et al., 2022).

Embora vários estudos tenham demonstrado reduções nos títulos de neutralização dos anticorpos induzidos por vacinas contra as variantes do SARS-CoV-2, os imunizantes disponíveis atualmente têm demonstrado excelentes resultados de eficácia e eficiência. Estes resultados estão refletidos principalmente na redução do número de hospitalizações e vidas perdidas pela Covid-19 (EMBORG et al., 2021; HAAS et al., 2021; HUANG et al., 2022; RASHEDI et al., 2022).

Ainda nesse sentido sobre a eficácia dos imunizantes, mais recentemente, o município de Botucatu, localizado no interior do estado de São Paulo no Brasil, foi palco de um importante estudo direcionado à eficácia da vacina ChAdOx1-nCov-19 - AZD1222 (AstraZeneca) sobre as variantes do SARS-CoV-2. Esse estudo está sendo conduzido em parceria da Unesp (Universidade Estadual Paulista) com a Universidade de Oxford, com o Ministério da Saúde, com a Prefeitura de Botucatu, com a Fiocruz (Fundação Oswaldo Cruz) e também com o Instituto Bill e Melinda Gates. O estudo se iniciou com a vacinação em massa no dia 16 de maio de 2021. A meta era vacinar 60 mil moradores, entretanto foram alcançadas mais de 66 mil pessoas, incluindo os estudantes da Unesp que residem na cidade (PREFEITURA DE BOTUCATU, 2021a).

Até a data inicial da aplicação da vacina, Botucatu acumulava mais de 12 mil casos confirmados e 207 óbitos (PREFEITURA DE BOTUCATU, 2022b). Números bastante expressivos, onde o número de casos positivos já representava aproximadamente 8% do total de habitantes. Nesse estudo, foram incluídos os moradores de Botucatu com idade entre 18 e 60 anos. As pessoas com mais de 60 anos de idade já tinham iniciado algum tipo de esquema vacinal e continuaram sendo imunizadas de acordo com o Plano Nacional de Imunização. As pessoas com idade inferior a 18 anos ainda não possuíam aprovação para receberem as doses de vacinas, entretanto, foram incluídas no Plano Nacional de Imunização por volta de junho de 2021 (MINISTÉRIO DA SAÚDE, 2022).

Esse cenário de surgimento de novas variantes do SARS-CoV-2 resistentes juntamente com o estudo da efetividade de uma vacina sobre essas variantes no município de Botucatu, deixou algumas questões importantes a serem discutidas. Como por exemplo, qual ou quais as variantes do SARS-CoV-2 predominavam no contexto pré-vacinação? A predominância dessas variantes pode estar associada a diferentes grupos de pessoas de acordo com sexo, idade e presença de comorbidades? As variantes ou mutações no genoma viral se correlacionam com a gravidade da Covid-19?

O atual contexto da emergência de novas variantes do SARS-CoV-2 e o agravamento da pandemia da Covid-19 torna urgente a necessidade de esforços de vigilância genômica robustos. O sequenciamento genômico, o compartilhamento em tempo real desses dados e o estudo das alterações no genoma viral, além de orientar o entendimento do percurso evolutivo do SARS-CoV-2, podem subsidiar as tomadas de decisões sobre o manejo da Covid-19. Esses estudos podem, por exemplo, contribuir com os esforços aprimorados de cobertura vacinal, com a distribuição de anticorpos monoclonais e outras terapias, com as decisões públicas sobre as medidas restritivas, etc.

Pensando nessas questões, o objetivo central do nosso estudo se concentra na avaliação da dispersão temporal e epidemiológica das variantes do SARS-CoV-2 no cenário de pré-vacinação em massa no município de Botucatu-SP.

## 1.1 Objetivos

### 1.1.1 Geral

Avaliar a dispersão temporal e epidemiológica das variantes do SARS-CoV-2 no período de pré-vacinação em massa no município de Botucatu-SP.

### 1.1.2 Específicos

Analisar a prevalência das variantes do SARS-CoV-2 nas quatro Semanas Epidemiológicas (SE 16, 17, 18 e 19 de 2021) que antecedem o dia de vacinação em massa em Botucatu-SP.

Investigar a prevalência das linhagens do SARS-CoV-2 em pacientes que convivem com alguma comorbidade e em pacientes pertencentes a diferentes faixas etárias e gêneros.

Correlacionar as variantes do SARS-CoV-2 com os diferentes espectros clínicos da Covid-19 – casos sintomáticos e assintomáticos, leves ou moderados, graves e óbitos.

## 2 METODOLOGIA

### 2.1 Seleção das amostras para o estudo

Por contar com amostras biológicas e dados clínicos e epidemiológicos de seres humanos, o presente projeto foi submetido e aprovado pelo Comitê de Ética em Pesquisa (CEP) da Faculdade de Medicina de Botucatu-SP (CAAE: 43218921.6.0000.5411).

Como critério de inclusão das amostras, foi feita uma lista considerando todas as amostras positivas coletadas em Botucatu-SP nas 4 Semanas Epidemiológicas (SE 16, 17, 18 e 19) (SINAN, 2021) que antecederam o dia da vacinação massiva, que ocorreu em 16 de maio de 2021. Essas SE datam do dia 18 de abril até o dia 15 de maio de 2021. Essas amostras são de pacientes sintomáticos e assintomáticos de todas as idades que procuraram algum sistema de saúde de Botucatu, tanto particular quanto público, e testaram positivo para SARS-CoV-2 por RT-qPCR (Transcrição Reversa Quantitativa). Além disso, consideramos as amostras com valores de  $C_T$  de até 30 elegíveis para o sequenciamento. A partir dessa lista de amostras positivas, selecionamos 400 amostras de maneira randomizada utilizando o Microsoft Office Excel, sendo 100 amostras por SE.

As informações clínicas dos participantes foram extraídas dos relatórios de notificação de casos suspeitos de Covid-19 (E-Sus) e de internação por Síndrome Respiratória Aguda Grave (Sivep-Gripe). Para recuperar a situação vacinal dos participantes foi utilizada a base dados VACIVIDA.

### 2.2 Obtenção das amostras positivas para SARS-CoV-2

As amostras contendo o RNA total foram gentilmente cedidas pelo Laboratório de Biologia Molecular localizado no Hospital das Clínicas da Faculdade de Medicina de Botucatu (HCFMB). Essas amostras são de secreção coletada em região de naso/orofaringe com *swabs* estéreis. Após a coleta, as amostras foram armazenadas em freezers a -80 °C até a extração do RNA total.

### 2.3 Extração de RNA e confirmação da presença do SARS-CoV-2 por RT-qPCR

A extração do RNA total das amostras foi realizada utilizando uma metodologia de Imobilização Reversível em Fase Sólida (Solid Phase Reversible Immobilization - SPRI)

desenvolvida pela própria equipe da rotina de diagnóstico do hospital. Uma alíquota desse RNA extraído foi usada para a detecção do SARS-CoV-2 por RT-qPCR e outra alíquota foi armazenada a -80 °C para o sequenciamento.

A confirmação da presença do vírus nas amostras foi feita por RT-qPCR utilizando o *kit Allplex™ SARS-CoV-2 Assay* (Seegene). Esse ensaio amplifica os genes RdRP, S e N específicos para o SARS-CoV-2 e o gene E para todos os Sarbecovírus, incluindo SARS-CoV-2.

## 2.4 Preparo das bibliotecas e Sequenciamento do Genoma do SARS-CoV-2

O sequenciamento completo do genoma do SARS-CoV-2 foi realizado em duas plataformas de acordo com a disponibilidade de reagentes no laboratório. As plataformas utilizadas foram Illumina e Ion Torrent da Thermo Fisher Scientific.

### 2.4.1 Preparo das bibliotecas e sequenciamento Illumina

Para o sequenciamento do genoma do SARS-CoV-2 utilizando a plataforma da Illumina, foi usado o *kit COVIDSeq Test* (Illumina®) para a construção das bibliotecas de acordo com as instruções do fabricante. As reações foram executadas em placas de PCR de 96 poços.

Resumidamente, o cDNA foi sintetizado a partir do RNA extraído utilizando hexâmetros aleatórios. Posteriormente o genoma viral foi amplificado utilizando dois *pools* de *primers* específicos para SARS-CoV-2 em duas reações de PCR *multiplex* separadas. Esses *pools* também contêm *primers* para amplificar genes humanos para o controle da expressão endógena. O produto da PCR foi marcado com adaptadores e com *Indexes* IDT para *Illumina Nextera UD Indexes* (Set A, B, C e D). O enriquecimento e a purificação das bibliotecas foram seguidos de acordo com as instruções do fabricante.

As 96 amostras foram agrupadas em um único *pool*, que foi purificado e quantificado posteriormente. A quantificação foi feita utilizando o *kit Qubit dsDNA High Sensitivity assay* em um fluorômetro *Qubit* (Invitrogen Inc, Carlsbad, CA, USA). O *pool* de bibliotecas foi normalizado a 4nM e denaturado com 5 µL de NaOH a 0,2 N. A biblioteca foi diluída a 7 pM e combinada com 1% de controle *PhiX* (*PhiX Control v3*, Illumina Inc, San Diego, CA, USA). As amostras foram sequenciadas em um MiSeq utilizando uma *Flow Cell* v2 de 300 ciclos (*MiSeq Reagent Kit v2 - 300-cycles*).

#### 2.4.1.1 Fluxo de bioinformática - Illumina

As leituras foram desmultiplexadas e os arquivos *FASTQ* foram gerados no próprio sequenciador. As sequências do vírus foram alinhadas com o Genoma de Referência *Wuhan-Hu-1 NCBI* (Número de Acesso: MN908947.3) por meio do *pipeline* de um aplicativo *online* da Illumina chamado *DRAGEN COVID Lineage 3.5.5*. Este *pipeline* também retorna alguns resultados de qualidade, como cobertura e profundidade de sequenciamento do genoma viral. As linhagens do SARS-CoV-2 presentes nas amostras foram determinadas utilizando o *software online Pangolin COVID-19 Lineage Assigner*, versão 4.0.6 (<http://pangolin.cog-uk.io/>, acessado em 25 de maio de 2022). A ferramenta *online Nextclade* versão 1.14.1 (<https://clades.nextstrain.org/>, acessada em 25 de maio de 2022) também foi utilizada para atribuição de clado e para a verificação de qualidade das sequências. As amostras com um Status geral do Controle de Qualidade “bom” foram incluídas nas análises posteriores, enquanto as amostras com status “ruim ou medíocre” foram desconsideradas.

#### 2.4.2 Preparo das bibliotecas e sequenciamento ION Torrent

O RNA total das amostras foi submetido a uma reação de transcrição reversa para a síntese do cDNA utilizando o *kit SuperScript VILO* (Thermo Fisher Scientific®) de acordo com as instruções do fabricante. O *kit Ion AmpliSeq SARS-CoV-2* (Thermo Fisher Scientific®) foi utilizado para o preparo das bibliotecas.

Após a síntese de cDNA, foi feita uma amplificação do genoma viral em duas reações de PCR multiplex separadas, utilizando dois *pools* de primers específicos para o SARS-CoV-2 (*Ion AmpliSeq SARS-CoV-2 Insight Research Assay GS, manual* - Thermo Fisher Scientific®). Esses *pools* também contêm *primers* para amplificar genes humanos para o controle da expressão endógena. Nessa reação são produzidos *amplicons* de 237 fragmentos do genoma do SARS-CoV-2 e de 5 genes humanos. Os *amplicons* do genoma viral variam de 125 a 275 pares de bases (pb) de tamanho e cobrem >99% do genoma do vírus.

Após isso, os *amplicons* foram parcialmente digeridos, marcados com adaptadores *barcodes* (*ION Xpress* da Thermo Fisher Scientific®) e purificados com *AMPure XP Beads* (Beckman Coulter®). Posteriormente, as bibliotecas foram amplificadas, purificadas e normalizadas a ~100 pM, utilizando o *Ion Library Equalizer* (Thermo Fisher Scientific®). *Pools* de 48 amostras foram feitos e posteriormente quantificados por qPCR utilizando o *kit Ion Library Taq Man Quantification* em um *QuantStudio* (Thermo Fisher Scientific®). Os

*pools* de bibliotecas foram diluídos a 50 pM e carregados em chips *Ion 540* utilizando um *Ion Chef* (Thermo Fisher Scientific®). As amostras foram sequenciadas em um sequenciador *Ion S5* (Thermo Fisher Scientific®).

#### 2.4.2.1 Fluxo de bioinformática – *Ion Torrent*

As leituras produzidas foram alinhadas ao Genoma de Referência *Wuhan-Hu-1 NCBI* (Número de Acesso: MN908947.3) utilizando o *pipeline online Torrent Suite v. 5.12.1* da Thermo Fisher Scientific®. Este *pipeline* também retorna os resultados de qualidade, como cobertura e profundidade de sequenciamento do genoma. As linhagens do SARS-CoV-2 presentes nas amostras foram determinadas utilizando o *software online Pangolin COVID-19 Lineage Assigner*, versão 4.0.6 (<http://pangolin.cog-uk.io/>, acessado em 25 de maio de 2022). A ferramenta *online Nextclade* versão 1.14.1 (<https://clades.nextstrain.org/>, acessada em 25 de maio de 2022) também foi utilizada para a atribuição de clado e para a verificação de qualidade das sequências. As amostras com um Status geral do Controle de Qualidade “bom” foram incluídas nas análises posteriores, enquanto as amostras com status “ruim ou medíocre” foram desconsideradas.

### 2.5 Análise estatísticas dos Resultados

Como dito anteriormente a determinação das variantes do SARS-CoV-2 e a análise de qualidade dos dados foram realizadas nos softwares *online* PANGO e *NextClade*. Para a tabulação e randomização das amostras foi utilizado o aplicativo Excel do pacote de aplicativos *Microsoft Office Professional Plus 2016*. O mesmo aplicativo foi utilizado para calcular as frequências em porcentagem das variantes do SARS-CoV-2. Os resultados foram plotados em gráficos produzidos no *GraphPad Prism 8.0.2* e no Excel.

Para comparar a frequência das linhagens do SARS-CoV-2 com as variáveis selecionadas para este estudo, foi feito o Teste da associação por qui-quadrado (SATORRA; BENTLER, 2001). Após isso foi feito o Teste de comparação das proporções por qui-quadrado para analisar as linhagens mais predominantes encontradas neste estudo. Para a variável utilizando os valores de  $C_T$  foi utilizado o anova para a comparação das médias. Enquanto, para analisar a soma das comorbidades, foi utilizado o teste de *Poisson* (FROME; KUTNER; BEAUCHAMP, 1973). Essas análises foram executadas no *software SAS* para



*windows*, v.9.4. e o no R v.3.6.1. As análises estatísticas foram realizadas em conjunto com o Escritório de Apoio à Pesquisa (EAP) da UNESP-FMB.

## 2.6 Análises Filogenéticas

Foi utilizado o fluxo de trabalho *Nextstrain v3.0.3* SARS-CoV-2 (HADFIELD et al., 2018) para as análises filogenéticas. Para isso foram utilizadas as 371 sequências das 4 SE (16, 17, 18 e 19 de 2021) antecedentes a vacinação em massa de Botucatu geradas neste estudo. Um conjunto de dados global representativo foi recuperado do GISAID (2.725 sequências, baixado de diversas regiões globais, com amplitude de data de coleta de 13 de janeiro de 2020 até 22 de julho de 2022 - ID das amostras GISAID (*Global Initiative on Sharing All Influenza Data*) – incluindo as sequências de 2019 usadas como referência) (<https://www.gisaid.org/>) que foi usado como *background* para as nossas análises. A árvore filogenética foi visualizada no *auspice.us 0.12.0* (Auspice 2.37.3) (<https://auspice.us/>).

### 3 RESULTADOS

#### 3.1 Amostras sequenciadas e qualidade do sequenciamento

Selecionamos de maneira randomizada um total de 400 amostras positivas para SARS-CoV-2 que foram coletadas entre as datas 18/04/2021 e 15/05/2021 no HCFMB de pessoas que residem em Botucatu. Esse período equivale as quatro Semanas Epidemiológicas (SE 16, 17, 18 e 19 de 2021) que precederam o dia de vacinação em massa que ocorreu no município. Foram Selecionadas 100 amostras de cada SE.

Duas plataformas foram usadas para sequenciar o genoma do vírus, sendo a plataforma *Illumina* e *Ion Torrent*. Do total de 400 amostras submetidas ao sequenciamento, 371 (92,75%) leituras passaram nos controles de qualidade das respectivas plataformas de sequenciamento e da ferramenta *Nextclade*. Na tabela 1 encontra-se a relação das datas das SE selecionadas no estudo, bem como o total de casos positivos registrados na cidade de Botucatu nesse período e também o número de sequências de cada plataforma que passaram nos controles de qualidade.

**Tabela 1: Relação de amostras positivas e sequenciadas nas SE selecionadas do estudo. Encontram-se também a quantidade de genomas que passaram em todos os testes de qualidade e as plataformas de sequenciamento**

|                                                                  | SE 16<br>18/04/2021 a<br>24/04/2021 | SE 17<br>25/04/2021 a<br>01/05/2021 | SE 18<br>02/05/2021 a<br>08/05/2021 | SE 19<br>09/05/2021 a<br>15/05/2021 |
|------------------------------------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Total de casos positivos em Botucatu                             | 373                                 | 379                                 | 481                                 | 608                                 |
| Número de amostras selecionadas                                  | 100                                 | 100                                 | 100                                 | 100                                 |
| Percentual amostrado                                             | 27%                                 | 26%                                 | 21%                                 | 16%                                 |
| Quantidade de sequencias que passaram nos controles de qualidade | 92                                  | 93                                  | 90                                  | 96                                  |
| Quantidade de sequências produzidas na plataforma da Illumina    | 7                                   | 67                                  | 25                                  | 0                                   |
| Quantidade de sequências produzidas na plataforma Ion Torrent    | 85                                  | 26                                  | 65                                  | 96                                  |

Tabela 1: Relação de amostras positivas e sequenciadas nas SE selecionadas do estudo. Encontram-se também a quantidade de genomas que passaram em todos os testes de qualidade e as plataformas de sequenciamento. Fonte: Elaborada pelo próprio autor.

Das 400 amostras submetidas no sequenciamento, 7 falharam na preparação das bibliotecas ou no sequenciamento por razões desconhecidas. Essas amostras possuíam valores de PCR em tempo real ( $C_T$ ) variando de 14,53 a 28,83. Além disso, 11 amostras obtiveram o *Status* geral do Controle de Qualidade “ruim” e outras 11 amostras obtiveram o *status* “mediocre” na ferramenta *online Nextclade*. Essas amostras foram desconsideradas nas análises posteriores. Os valores de  $C_T$  das amostras que não passaram nos controles de qualidade do Nextclade variavam de 15,6 a 29,82.

Para esse estudo, consideramos as 371 sequências de alta qualidade com pelo menos 98% de cobertura do genoma de referência do SARS-CoV-2 (GenBank MN985325) e um mínimo de profundidade de 50x (PLITNICK et al., 2021). No presente estudo, o conjunto amostral alcançou uma média de profundidade de 6.290x. Além disso, esses genomas passaram pelos controles de qualidade dos softwares *online PANGO* e *NextClade* (*Status* geral do Controle de Qualidade “bom”).

### 3.2 Descrição dos dados demográficos, epidemiológicos e clínicos dos participantes do estudo

Entre os 371 participantes incluídos neste estudo, 53,6% eram pessoas do sexo feminino e 46,4% eram do sexo masculino. A idade média dos participantes foi de 35,4 anos (Desvio Padrão = 19,16), sendo a idade mínima 3 meses e a idade máxima 90 anos. Os valores de PCR em tempo real ( $C_T$ ) para a presença do SARS-CoV-2 nessas amostras, estavam na média de 20,3 sendo o menor valor 11 e o maior valor 29,9.

Um total de 61 participantes neste estudo convivia com algum tipo de comorbidade, enquanto 292 pessoas não portavam nenhuma comorbidade e 18 pessoas não continham informações sobre essa variável nos bancos de dados. As comorbidades incluídas neste estudo, bem como o número de participantes que convivem com essas doenças, estão expressos na tabela 2.

| Tabela 2: Relação das comorbidades importantes para Covid-19 encontradas neste estudo e a quantidade de participantes que convivem com essas doenças. |                             |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|
| Comorbidades                                                                                                                                          | Quantidade de participantes |
| Artrite Reumatoide                                                                                                                                    | 1                           |
| Diabetes                                                                                                                                              | 18                          |
| Doenças Cardiovasculares                                                                                                                              | 30                          |

|                                                                        |    |
|------------------------------------------------------------------------|----|
| Doenças Renais                                                         | 2  |
| Doenças Respiratórias Crônicas Descompensadas (DRCD)                   | 16 |
| Ex-etilista                                                            | 1  |
| Ex-tabagista                                                           | 3  |
| Gestante                                                               | 1  |
| Imunossupressão                                                        | 4  |
| Neuropatia                                                             | 1  |
| Pneumopatia                                                            | 1  |
| Portador de doenças cromossômicas ou estado de fragilidade imunológica | 1  |

Tabela 2: Relação das comorbidades importantes para Covid-19 encontradas neste estudo e a quantidade de participantes que convivem com essas doenças. Fonte: Elaborada pelo próprio autor.

Em relação às pessoas imunizadas, até a data do início do presente estudo, 48 pessoas tinham sido imunizadas, sendo 18 pessoas vacinadas com a primeira dose da AZD1222 (Astrazeneca) e em 2 pessoas já havia sido administrada a segunda dose deste mesmo imunizante. Outras 5 pessoas foram vacinadas com uma dose da CoronaVac (Sinovac) e 23 pessoas foram imunizadas até a segunda dose desta mesma vacina.

Também foram coletadas informações clínicas a respeito dos sintomas da Covid-19. Nos bancos de dados continham informações de 350 participantes, e destes, mais de 73% eram pacientes sintomáticos e aproximadamente 27% assintomáticos.

Quanto aos dados de hospitalizações, 21 pessoas (5,71%) precisaram oxigenoterapia, sendo que delas, 86% utilizaram suporte de oxigênio não invasivo, enquanto 14% dessas pessoas utilizaram a forma invasiva do suporte. Sobre o desfecho da doença, 4 pessoas (1%) foram a óbito. Essas pessoas que faleceram por Covid-19 necessitaram de suporte de oxigênio, sendo que 3 delas, utilizaram o suporte de oxigênio invasivo.

Em relação a classificação final da Covid-19, consideramos como Covid-19 leve os pacientes que tiveram o diagnóstico positivo e que não precisaram de internação. Para Covid-19 moderada, foram considerados os participantes positivos para SARS-CoV-2 que precisaram de internação, e que apesar de hospitalizados, não precisaram de internação em Unidade de Terapia Intensiva (UTI) e sobreviveram. Já para Covid-19 grave, foram considerados os pacientes positivos que precisaram de internação em UTI, e utilizaram algum suporte de oxigênio e que se curaram ou foram a óbito. No presente estudo, pouco mais de

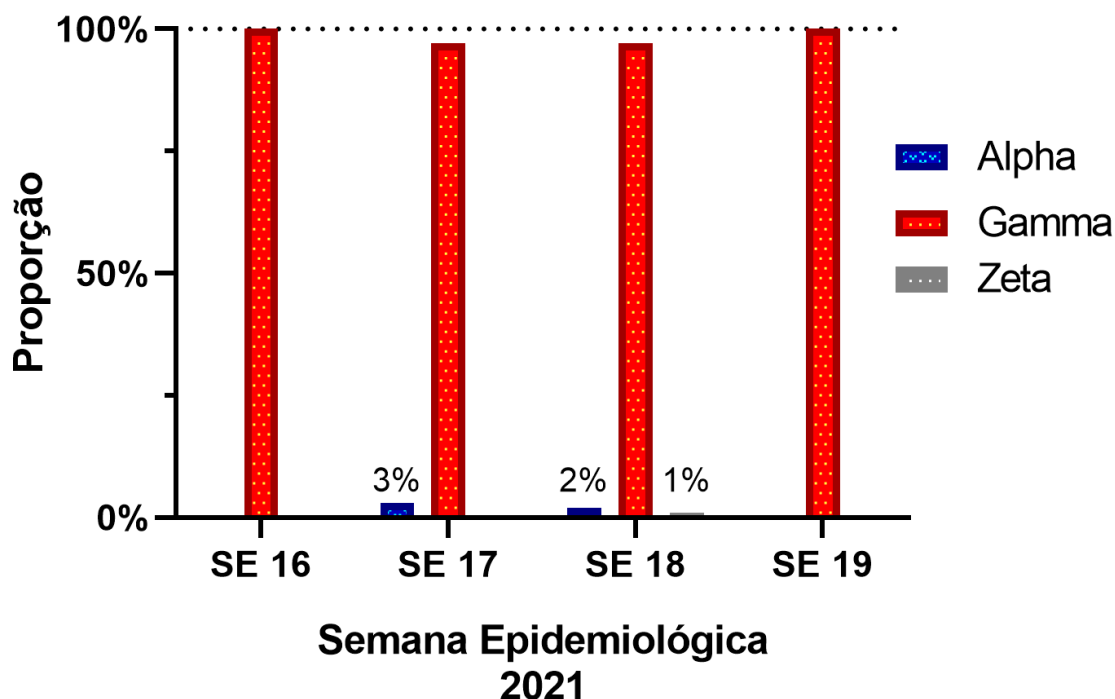
93% dos pacientes tiveram Covid-19 da forma leve, mais de 5% tiveram Covid-19 moderada e 1% teve Covid-19 grave.

### 3.3 Distribuição temporal das variantes do SARS-CoV-2 antes da vacinação em massa em Botucatu

Foi analisada a dispersão das variantes do SARS-CoV-2 através das SE 16, 17, 18 e 19 de 2021. Em todas as semanas foram encontradas apenas sequências de VOCs e VOI do vírus. Na SE 16, todas as amostras (92) correspondiam a VOC Gamma. Na SE 17, 96,8% das sequências (90) pertenciam a variante Gamma e 3,2% (3) a VOC Alpha. Na SE 18 foi registrado um total de 96,7% (87) de sequências de Gamma, 2,2% (2) de Alpha e apenas uma sequência da VOI Zeta. Na última semana do estudo, a SE 19, todas as amostras (96) foram da VOC Gamma.

Na Figura 1 estão representadas as frequências em porcentagem das variantes do SARS-CoV-2 encontradas nas 4 SE que antecederam o dia de vacinação em massa no município de Botucatu.

Figura 1: Distribuição temporal das variantes do SARS-CoV-2 nas SE 16, 17, 18 e 19 de 2021 no município de Botucatu-SP.



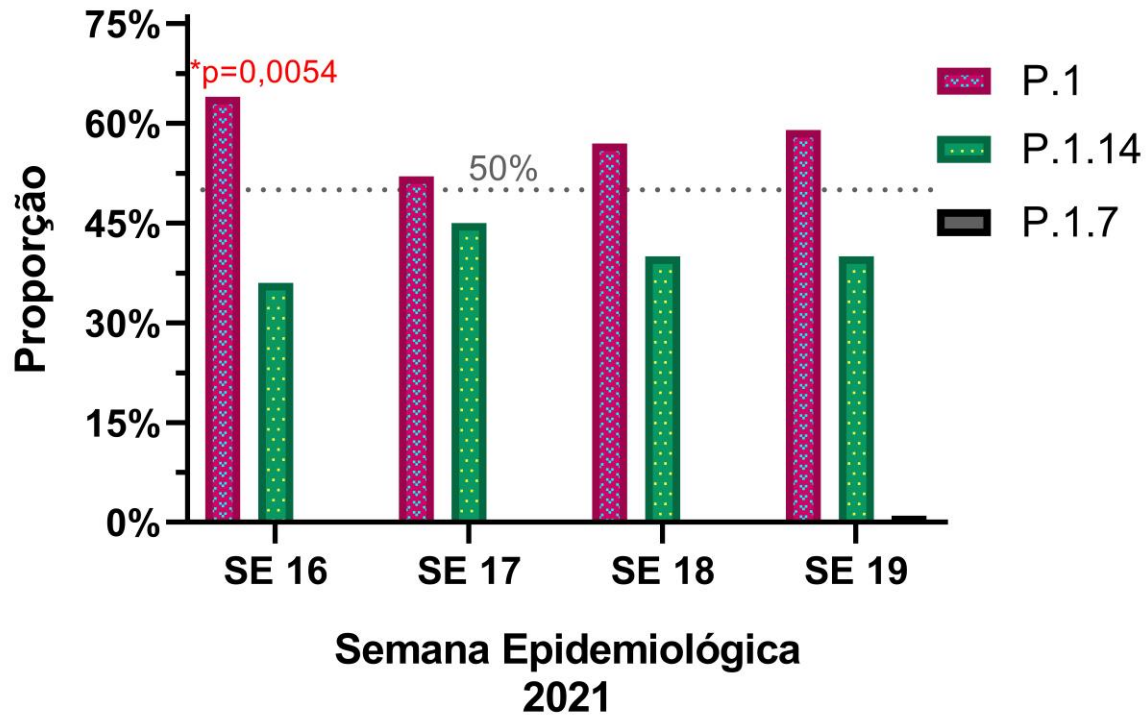
Fonte: Elaborada pelo próprio autor.

### **3.3.1 Distribuição das sublinhagens de Gamma do SARS-CoV-2 nas SE 16, 17, 18 e 19 de 2021**

Das 371 sequências produzidas nas SE selecionadas para o estudo, 365 correspondiam a variante Gamma do SARS-CoV-2, representando 98,4% do total de amostras sequenciadas. Mais análises sobre a dispersão dessa variante foram feitas considerando as suas sublinhagens. Pudemos observar que durante todo o estudo as linhagens P.1 e P.1.14 foram as mais frequentes.

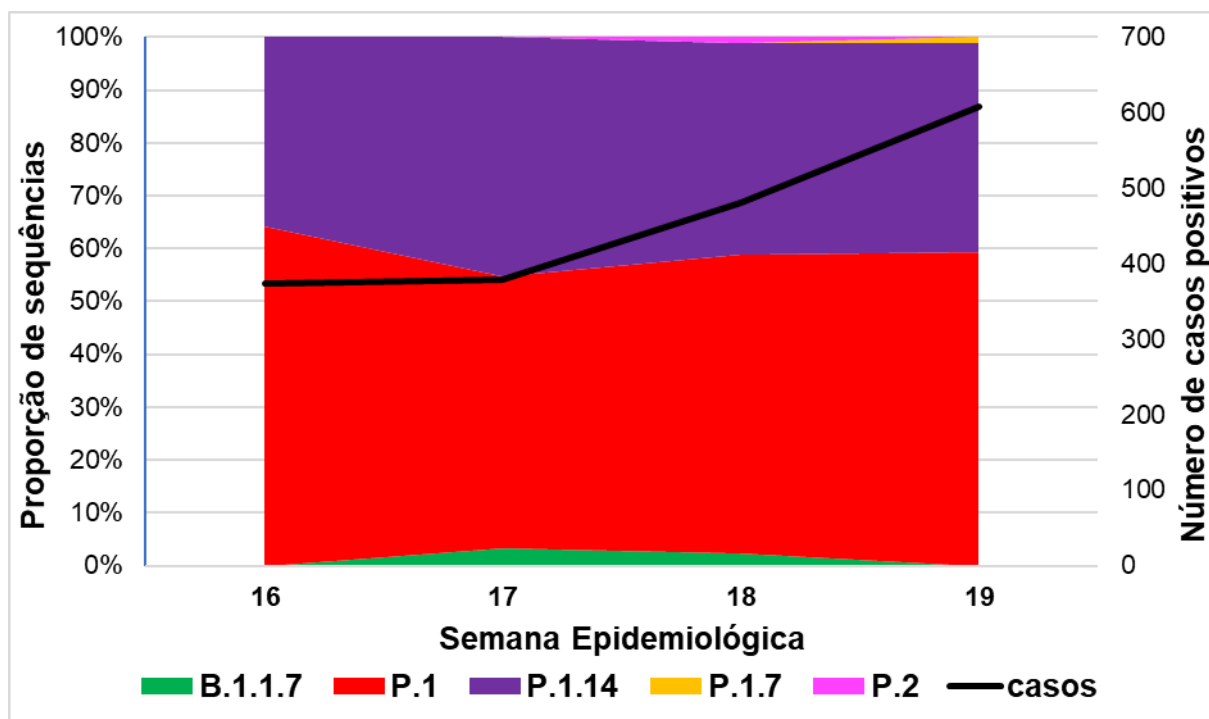
Na SE 16, a linhagem P.1 foi significativamente mais incidente em relação a sua sublinhagem P.1.14 ( $p=0,0054$ ). Do total de sequências produzidas nessa SE, 64,1% (59) e 35,9% (33) correspondiam às linhagens P.1 e P.1.14 respectivamente. Na SE 17, 51,6% (48) das amostras continham P.1, enquanto 45,2% (42) correspondiam a P.1.14. Já na SE 18 as frequências de 56,7% (51) e 40% (36) foram registradas respectivamente para as linhagens P.1 e P.1.14. Enquanto na SE19, 59,4% (57) das sequências de Gamma foram de P.1, 39,6% (38) de P.1.14 e 1% (1) de P.1.7. Não houve diferenças estatísticas nas proporções de P.1 em comparação com P.1.14 nas SE 17, 18 e 19 ( $P=0,575$ ;  $p=0,1155$  e  $P=0,0544$  respectivamente). Estes resultados estão expressos nas figuras 2 e 3, sendo que a figura 3 contém todas as linhagens encontradas e uma curva representando o total de casos positivos registrados na cidade de Botucatu em cada SE.

Figura 2: Distribuição das sublinhagens da VOC do SARS-CoV-2 Gamma nas SE 16, 17, 18 e 19 de 2021 no município de Botucatu-SP. O valor de p (\*) indica a significância estatística encontrada, indicando a maior proporção de P.1 quando comparada com P.1.14 na SE 16



Fonte: Elaborada pelo próprio autor.

Figura 3: Proporção das variantes do SARS-CoV-2 encontradas através das SE 16, 17, 18 e 19 de 2021 em relação ao total de casos positivos de cada SE, representados pela linha preta.



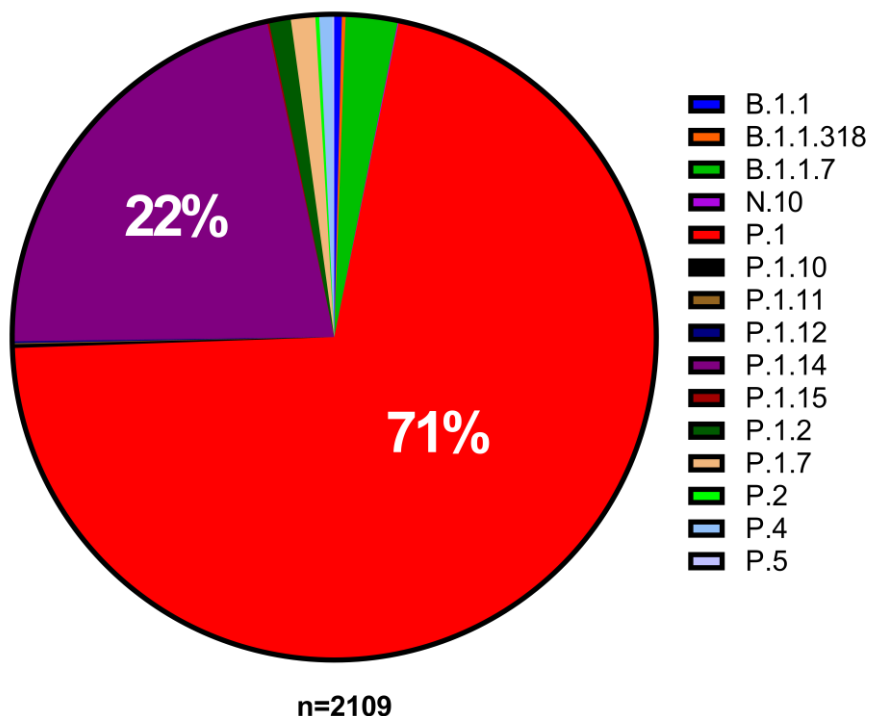
Fonte: Elaborada pelo próprio autor.

### 3.4 Comparação das frequências das linhagens do SARS-CoV-2 da cidade de Botucatu-SP com o estado de São Paulo

Além de calcular as frequências das linhagens encontradas em Botucatu-SP, também analisamos as frequências das variantes encontradas no estado de São Paulo durante o período do nosso estudo. As sequências e metadados das amostras estão depositadas no GISAID e também foram cedidas gentilmente pela Rede Butantan de Alerta Pandêmico de Variantes SARS-CoV-2 (Rede Butantan) do Estado de São Paulo. Das mais de 2 mil sequências produzidas no estado, a linhagem P.1 foi a mais frequente no estado de São Paulo durante o mesmo período deste estudo. A sublinhagem P.1.14 foi a segunda mais frequente. Os dados das variantes do SARS-CoV-2 durante as SE 16, 17, 18 e 19 de 2021 encontram-se na Figura 4.



Figura 4: Linhagens encontradas pela Rede Butantan de Alerta Pandêmico de Variantes SARS-CoV-2 durante a vigilância genômica do SARS-CoV-2 nas SE 16, 17, 18 e 19 de 2021.



Fonte: elaborada pelo próprio autor.

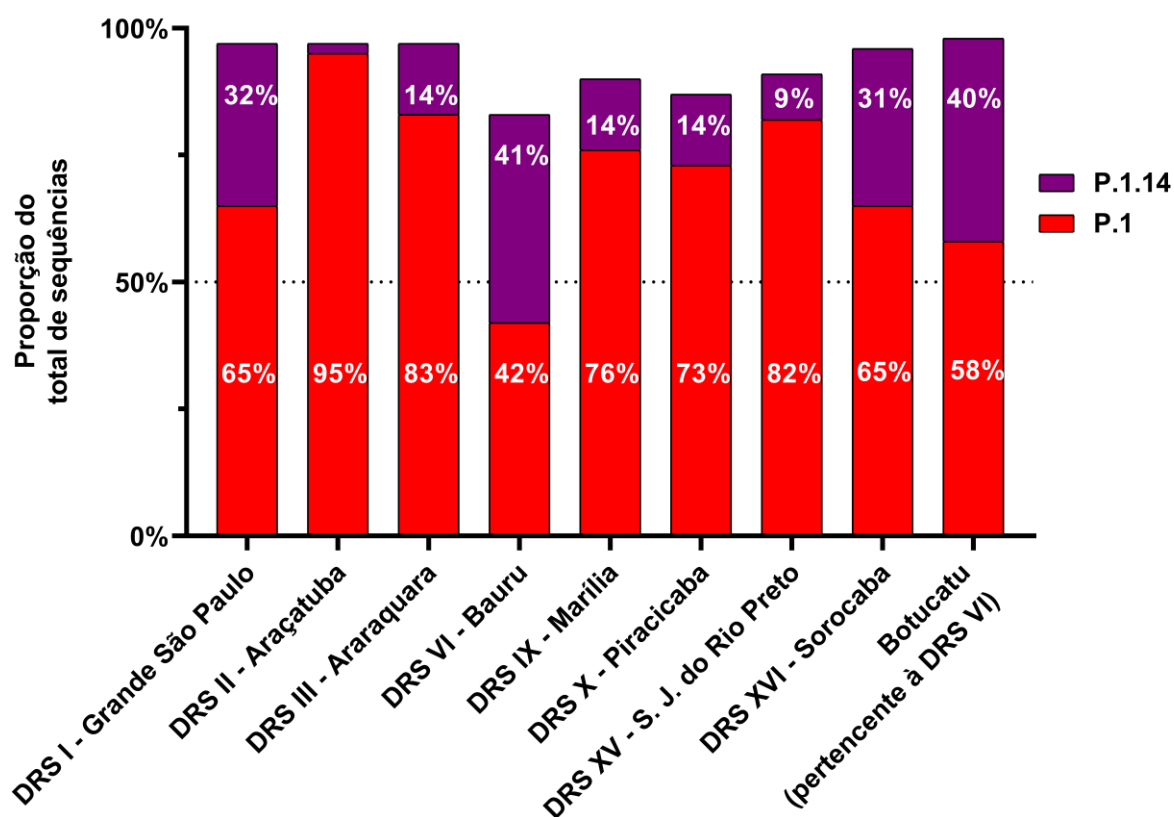
Também verificamos as frequências das linhagens P.1 e P.1.14 de acordo com a divisão administrativa da Secretaria de Estado da Saúde de São Paulo, que divide o estado por DRS (Departamentos Regionais de Saúde – Figura 5). Foram selecionadas as DRS que circulam geograficamente a DRS de Bauru, onde está localizada a cidade de Botucatu, e a DRS da capital do estado. Foi possível observar que a variante P.1 foi predominante na maioria das DRS. Na DRS IV onde se localiza o município de Botucatu, P.1 e P.1.14 obtiveram frequências praticamente idênticas. A figura 6 representa as frequências das linhagens P.1 e P.1.14 nas DRS que cercam a DRS de Bauru.

Figura 5: Departamentos Regionais de Saúde (DRS) do estado de São Paulo.



Fonte: mpsp.mp.br

Figura 6: Frequências das linhagens P.1 e P.1.14 nas DRS que circulam o município de Botucatu, incluindo a DRS da capital São Paulo.



Fonte: Elaborada pelo próprio autor.

### 3.5 Distribuição das variantes do SARS-CoV-2 de acordo com os gêneros

A análise de frequência das variantes do SARS-CoV-2 também foi feita levando em consideração o gênero dos participantes. Foi possível observar que a variante P.1 foi estatisticamente mais frequente do que a sua sublinhagem P.1.14 nos gêneros masculino e feminino ( $p=0,040$  e  $p<0.0001$  respectivamente).

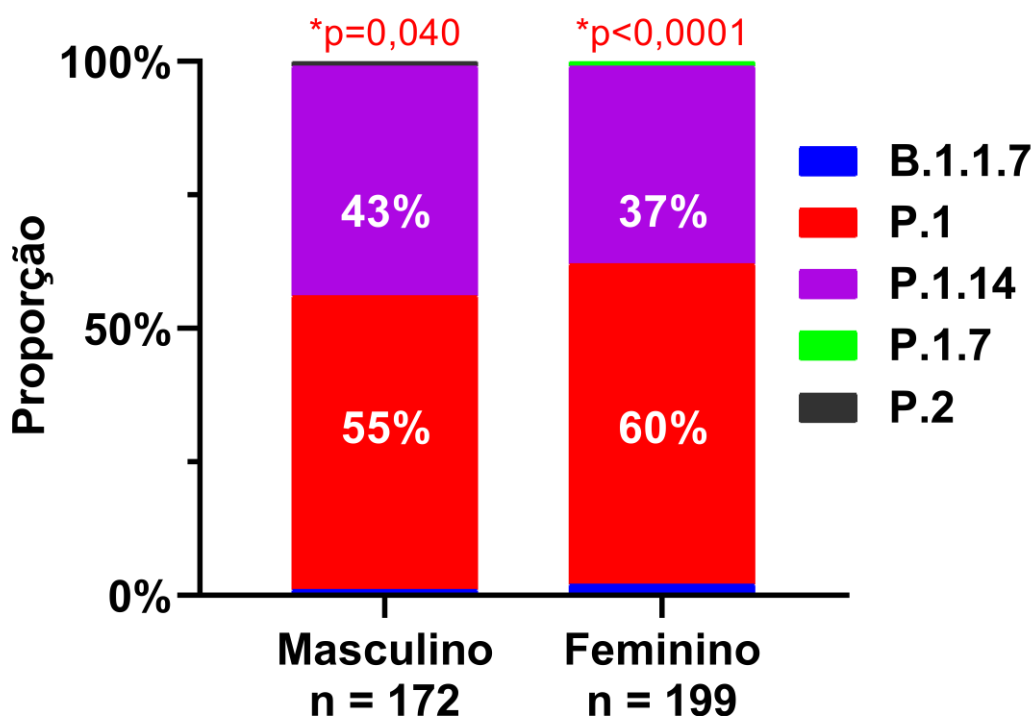
De todos os genomas do SARS-CoV-2 obtidos a partir de amostras do sexo feminino ( $n= 200$ ), 60,3% e 37,2% foram de P.1 e P.1.14 respectivamente. Em relação ao sexo masculino ( $n= 172$ ), 55,2% dos genomas sequenciados de amostras de pacientes desse gênero foram de P.1, enquanto 43,6% foram de P.1.14. Os dados completos de todas as variantes se encontram na tabela 3 e na figura 7.

**Tabela 3: Proporção da distribuição das variantes do SARS-CoV-2 do total de sequências produzidas em cada gênero.**

|         | <b>Feminino<br/>(n=200)</b> | <b>Masculino<br/>(n=172)</b> |
|---------|-----------------------------|------------------------------|
| B.1.1.7 | 2,0%                        | 0,6%                         |
| P.1     | 60,3%                       | 55,2%                        |
| P.1.14  | 37,2%                       | 43,6%                        |
| P.1.7   | 0,5%                        | 0,0%                         |
| P.2     | 0,0%                        | 0,6%                         |

Tabela 4: Proporção da distribuição das variantes do SARS-CoV-2 do total de sequências produzidas em cada gênero. Fonte: Elaborada pelo próprio autor.

Figura 7: Distribuição das variantes do SARS-CoV-2 de acordo com cada gênero. O valor de p (\*) indica a significância estatística encontrada nas frequências de P.1 quando comparada com a sublinhagem P.1.14.



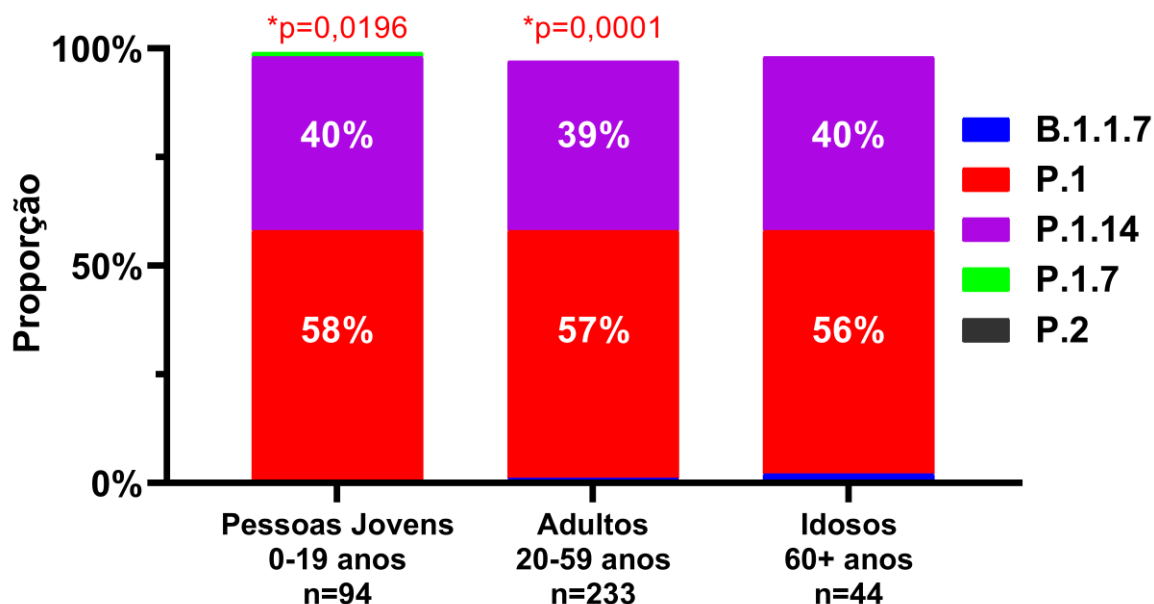
Fonte: Elaborada pelo próprio autor.

### 3.6 Distribuição das variantes do SARS-CoV-2 em diferentes faixas etárias

A distribuição das variantes do SARS-CoV-2 em função da faixa etária também foi analisada. Para isso, as sequências foram separadas em 3 grandes grupos. O grupo “Pessoas Jovens” é representado por pessoas de 0-19 anos de idade. O grupo “Adultos” contém pessoas com 20-59 anos de idade. Enquanto o grupo “Idosos” compreende aquelas pessoas com 60 anos ou mais. Os resultados estão expressos na figura 8.

Foi possível observar que P.1 foi significativamente mais predominante do que P.1.14 entre as pessoas mais jovens e adultas, representando 58,5% (p=0,0196) e 57,9% (p=0,0001) respectivamente. No grupo de pessoas idosas, quando comparadas as linhagens P.1 e P.1.14, não houve diferença estatística (p=0,2007) na frequência dessas variantes.

Figura 8: Frequência das variantes do SARS-CoV-2 distribuídas em diferentes grupos de acordo com as faixas etárias. O "n" indica o número total de amostras sequenciadas em cada grupo. O valor de p (\*) indica a significância estatística.



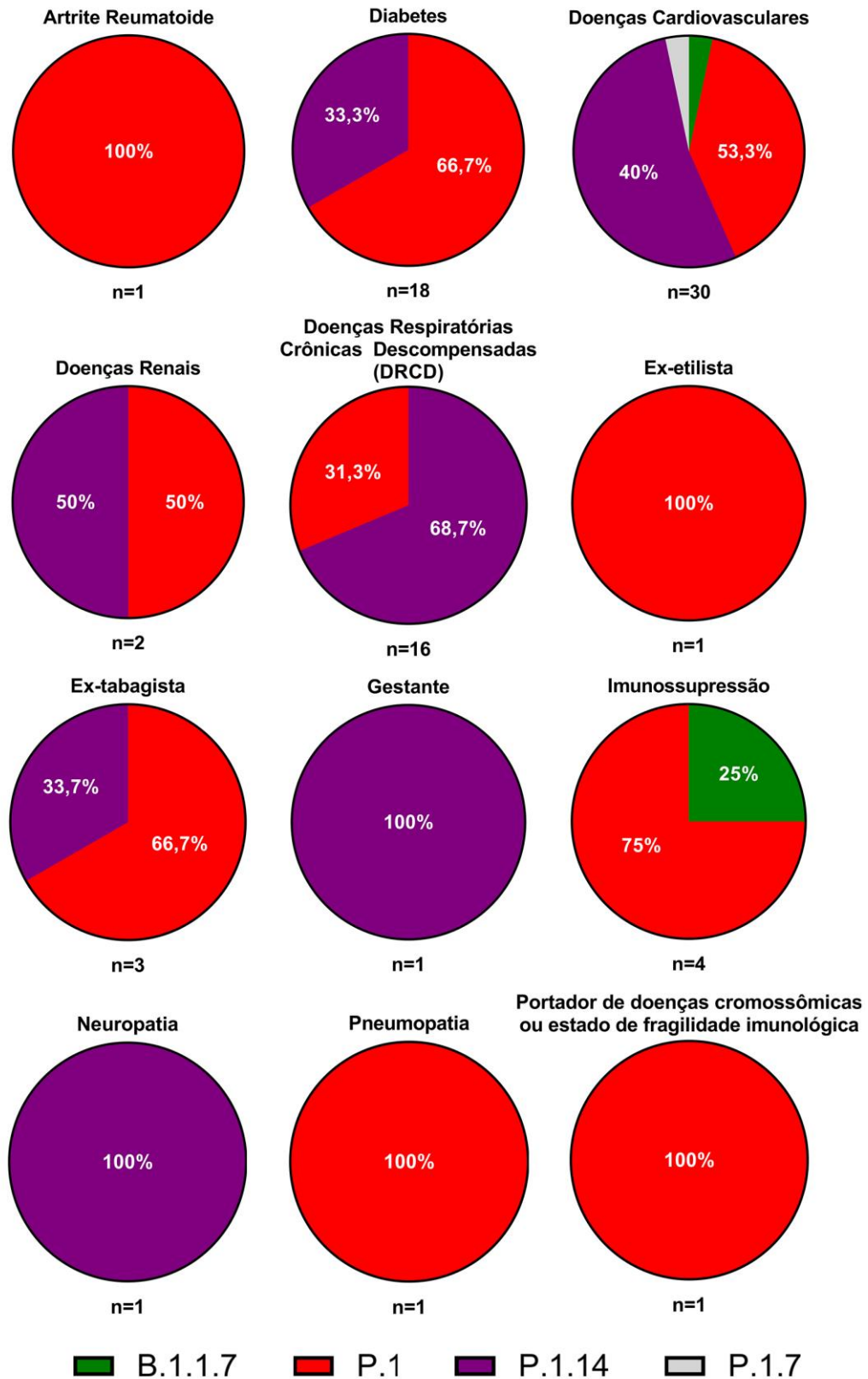
Fonte: Elaborada pelo próprio autor.

### 3.7 Associação das linhagens do SARS-CoV-2 com comorbidades

Foram encontradas 61 pessoas (16,4%) que convivem com alguma ou mais de uma doença de base ou comorbidade. Essas comorbidades e o número de participantes que convivem com essas doenças estão descritas na Tabela 2 deste manuscrito. Não foram encontradas associações significativas entre as linhagens e a presença destas comorbidades.

Além disso, é importante mencionar que não houve diferença estatística ao associar as linhagens do vírus com a presença de multicomorbidades ou policomorbidades ( $p=0,6221$ ). A figura 9 demonstra a frequência das linhagens do SARS-CoV-2 em cada comorbidade analisada.

Figura 9: Frequências das variantes do SARS-CoV-2 entre as pessoas com diferentes tipos de comorbidades. O "n" indica o número de participantes deste estudo que convivem com determinada comorbidade.



Fonte: Elaborada pelo próprio autor.

### 3.8 Associação das linhagens do SARS-CoV-2 com a situação clínica e vacinal dos participantes

Também buscamos investigar se algumas características clínicas dos participantes pudessem estar associadas com as linhagens P.1 e P.1.14. Foi observado que a variante P.1 foi significativamente ( $p=0,0343$ ) mais prevalente (54,1%) nos pacientes sintomáticos em comparação com a P.1.14 (44,4%). Por um outro lado, não houve diferença significativa nos valores de  $C_T$  quando comparamos P.1 com P.1.14.

Também foi encontrada diferença significativa nas frequências de P.1 e P.1.14 entre os pacientes que precisaram ou não de suporte de oxigênio (invasivo e não invasivo). Quanto ao quadro clínico da Covid-19, também não foi encontrada significância estatística na associação de P.1 em comparação com P.1.14 nos quadros leves, moderados e graves da Covid-19. Essas mesmas linhagens também não foram associadas estatisticamente aos casos que evoluíram a óbito.

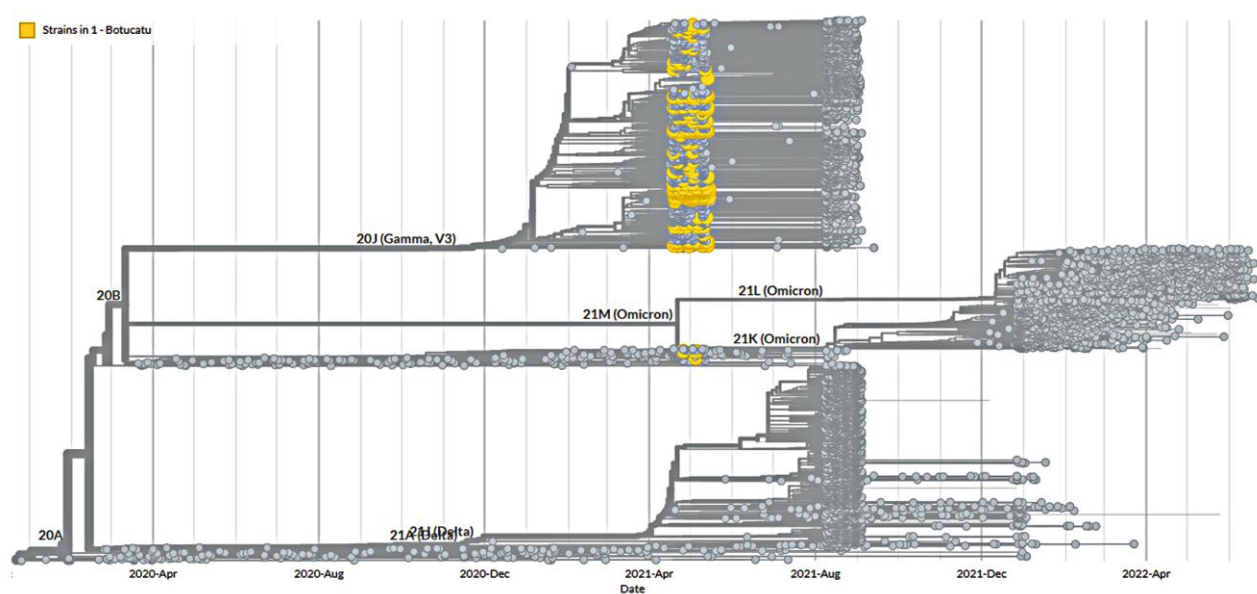
Em relação aos indivíduos vacinados, entre os 371 participantes, apenas 48 foram vacinados, sendo 28 participantes imunizados com CoronaVac e o restante com Astrazeneca. A linhagem P.1 foi mais prevalente que P.1.14 nas pessoas que tomaram CoronaVac ( $p=0,01589$ ), enquanto que não houve diferença significativa na frequência dessas variantes entre os participantes que tomaram a vacina da AstraZeneca. Quanto ao número de doses dessas vacinas, não foram encontradas diferenças nas frequências de P.1 e P.1.14 entre os indivíduos que tomaram uma ou duas doses.

### 3.9 Análises Filogenéticas

Foram utilizadas 371 sequências referentes as 4 SE (16, 17, 18 e 19 de 2021) do município de Botucatu-SP produzidas neste estudo, juntamente com um *background* de 2.725 sequências de diversas regiões do mundo. Essas amostras do *background* datam de 13 de janeiro de 2020 até 22 julho de 2022. Na figura 10 está a árvore filogenética gerada contendo as amostras de Botucatu (destacadas em amarelo) antes da vacinação em massa.



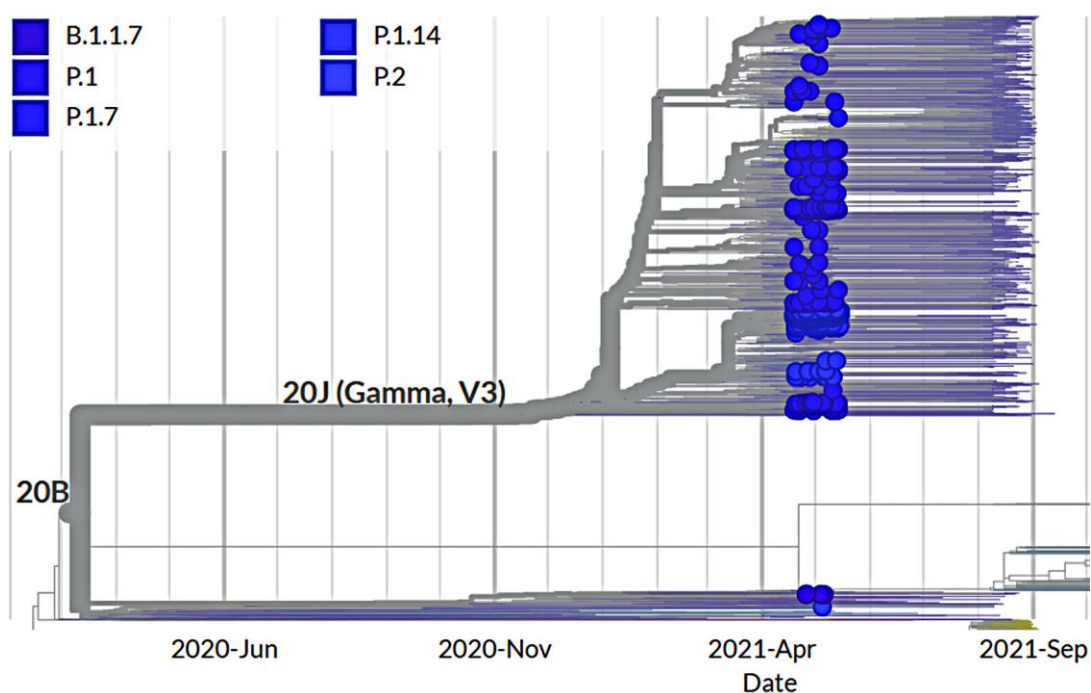
Figura 10: Árvore filogenética com as amostras de Botucatu-SP coletadas antes da vacinação em massa destacadas em amarelo.



Fonte: Elaborada pelo próprio autor.

Nesta árvore filogenética é possível observar que a grande maioria das amostras (98,4%) do presente estudo se agrupou no grande clado de Gamma, ou 20J. As figuras (Figuras 10, 11 e 12) demonstram esses achados. As outras amostras destacadas na parte inferior das figuras 10 e 11 estão representando as sequências de B.1.1.7 e P.2.

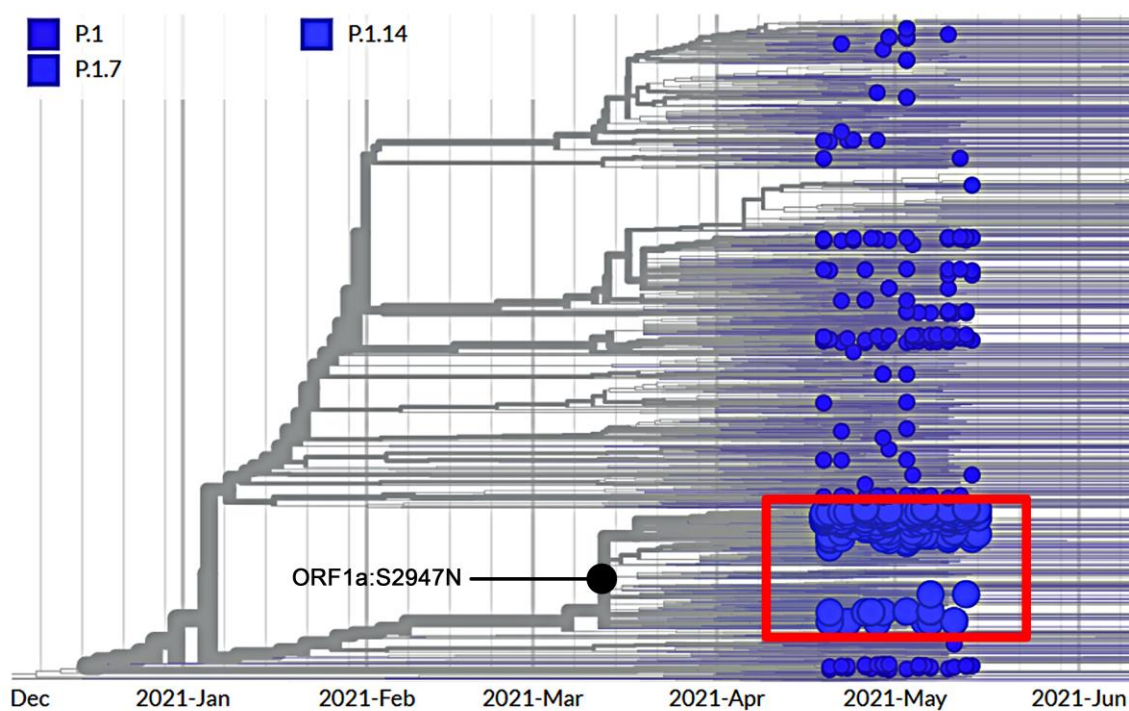
Figura 11: Árvore filogenética destacando o clado 20B contendo todas as 371 amostras do presente estudo realizado em Botucatu-SP nas SE 16, 17, 18 e 19 de 2021.



Fonte: Elaborada pelo próprio autor.

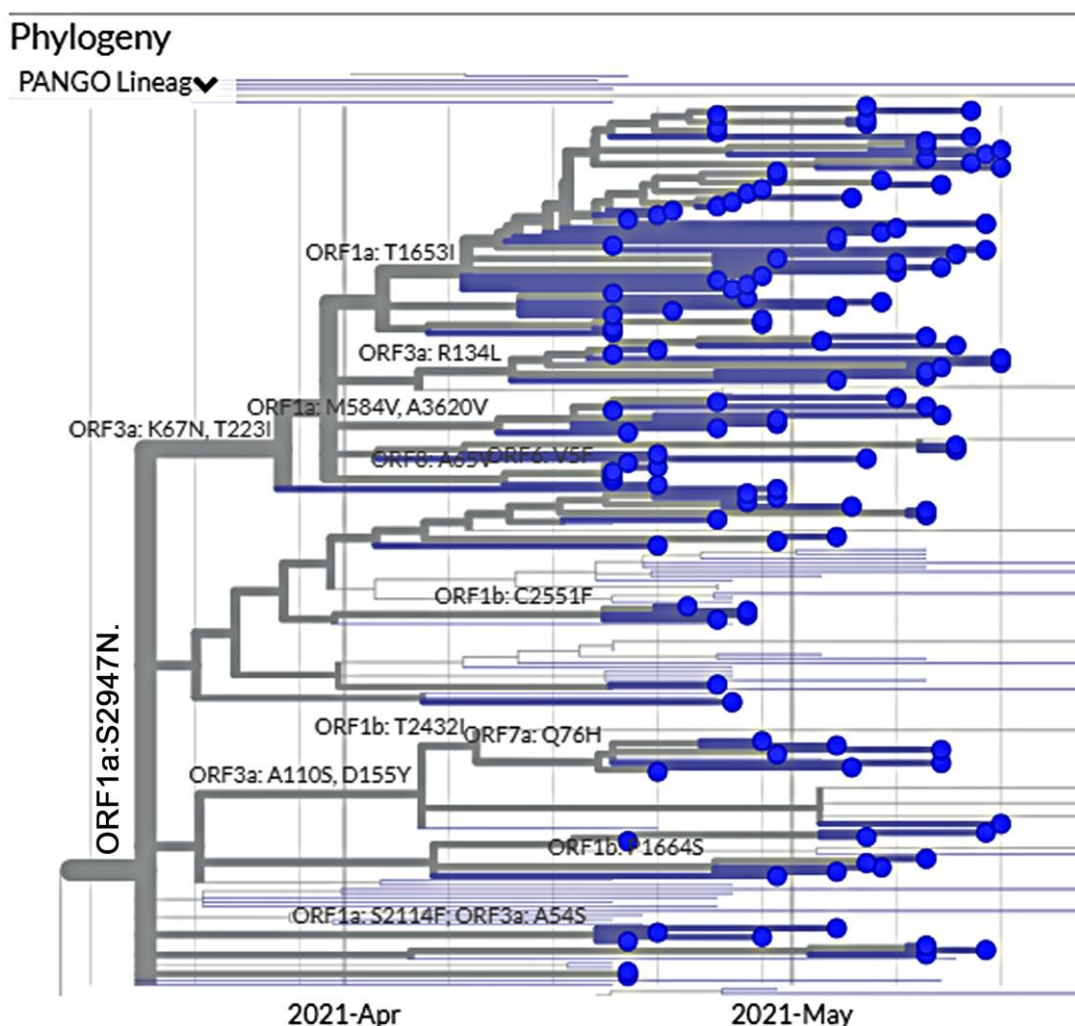
Dentre as amostras de Gamma, a linhagem P.1 se mostrou mais distribuída pelos ramos da árvore filogenética (Figura 12). Já a sua sublinhagem P.1.14 ficou mais agrupada em no grande ramo que possui a substituição em ORF1a:S2947N. Na figura 12, as amostras contendo P.1.14 estão destacadas dentro de um retângulo vermelho. Na figura 13 está árvore contendo a ORF1a:S2947N ampliada.

Figura 12: Árvore filogenética contendo apenas as linhagens de Gamma. Dentro do triângulo vermelho em destaque, estão as sequências da sublinhagem P.1.14 agrupadas no clado contendo a substituição em ORF1a:S2947N. As outras amostras são majoritariamente P.1.



Fonte: Elaborada pelo próprio autor.

Figura 13: Ramo da árvore da sublinhagem P.1.14 que contém a mutação em ORF1a:S2947N.



Fonte: Elaborada pelo próprio autor.

Ainda olhando para os ramos posteriores da ORF1a:S2947N, é possível observar um agrupamento de 9 sequências que se diferenciam pelas mutações em ORF1a: M584V e A3620V. Esse conjunto de duas mutações não constam em nenhuma sequência de P.1.14 no site [cov-spectrum.org/](https://cov-spectrum.org/) até o dia 05 de agosto de 2021. Essas amostras se encontram dentro do grande clado contendo as substituições em ORF3a: K67N e T223I e estão representadas na figura 13 e 14.



## 4 DISCUSSÃO

Entender como o SARS-CoV-2 persiste em ambientes onde uma alta proporção da população está desenvolvendo uma imunidade devido à ampla cobertura vacinal, é de suma importância para o entendimento do percurso evolutivo viral. Os estudos de prevalência e distribuição do SARS-CoV-2 podem impactar diretamente a saúde pública, fornecendo subsídios para o manejo das infecções pelo vírus e da Covid-19.

No município de Botucatu-SP, no dia 16 de maio de 2021 foi iniciada aplicação da primeira dose da vacina AZD1222 (Oxford-AstraZeneca) em todos os moradores adultos da cidade. A vacinação em massa ocorreu da mesma maneira que acontecem as votações no sistema eleitoral, ou seja, em apenas um dia mais de 66 mil pessoas foram imunizadas, comparecendo nas suas respectivas zonas eleitorais. Esse evento ocorreu devido ao estudo que foi realizado em Botucatu sobre a eficácia da vacina contra as variantes do SARS-CoV-2 (PREFEITURA DE BOTUCATU, 2021a).

Diante deste cenário buscamos entender como e quais as variantes do SARS-CoV-2 estavam circulando no município antes da vacinação em massa. Para isso, selecionamos, de maneira randomizada, um total de 400 amostras das SE 16, 17, 18 e 19 de 2021, sendo 100 amostra por SE. Utilizamos Sequenciamento de Nova Geração para produzir genomas completos e determinar as variantes do SARS-CoV-2 presentes nessas amostras.

Dessas 400 amostras, 371 sequências passaram em todos os controles de qualidade e seguiram para as análises posteriores. Nessas SE que antecederam a vacinação, a variante do SARS-CoV-2 predominante foi Gamma, representando 98,4% do total de sequências. Também encontramos sequências da VOC Alpha (1,3%) e apenas uma sequência da VOI Zeta (0,3%). Esses resultados, de certa forma, eram esperados, uma vez que a variante Gamma do SARS-CoV-2 predominava em grande parte do estado de SP e no Brasil no mesmo período do presente estudo (GIOVANETTI et al., 2021a). O crescimento da frequência de variantes do SARS-CoV-2 envolve vários fatores combinados que incluem a diversidade populacional dos hospedeiros e do vírus. O rápido crescimento e a prevalência da VOC Gamma podem ser explicados pelo aumento da transmissibilidade e um moderado potencial escape imunológico em relação às infecções causadas por outras variantes anteriores (CAMPBELL et al., 2021; CHAKRABORTY et al., 2021; COUTINHO et al., 2021; GIDARI et al., 2021).

A variante Gamma, linhagem P.1, é caracterizada por uma constelação de 24 mutações não sinônimas, incluindo 12 mutações apenas na proteína spike (NAVECA et al.,

2021a; covariants.org/variants/20J.Gamma.V3). A variante Gamma foi registrada pela primeira vez em 06 de janeiro de 2021 no Japão, a partir de viajantes que retornaram da capital Manaus localizada no Estado do Amazonas (NIID, 2021a). Entretanto, Gamma/P.1 foi encontrada em amostras que datam antes da sua descoberta (aproximadamente 2 meses) em vários estados brasileiros (LAMARCA et al., 2021). A VOC Gamma predominou por alguns meses em Manaus, provavelmente causando uma das maiores ondas de casos graves e óbitos na cidade, colapsando o sistema de saúde com uma crise de falta de respiradores e oxigênio (NAVECA et al., 2021a; NAVECA et al., 2021b). Posteriormente a variante Gamma se disseminou pelo restante do Brasil, se tornando predominante na maior parte do país até surgir a outra onda de casos provocados pela VOC Delta, por volta de julho de 2021 (GIOVANETTI et al., 2021a).

Wink et al., 2022, por exemplo, registraram a variante P.1 em mais de 97% das 126 amostras analisadas. Esse estudo foi realizado em um estado do sul do Brasil e essas sequências foram encontradas entre janeiro e maio de 2021. Em um outro estudo feito nesse mesmo estado por volta dessa mesma data, a variante Gamma também foi predominante, constando em 67% do total de amostras sequenciadas, seguida pelas linhagens P.2 e de sua ancestral B.1.1.28 (DEMOLINER et al., 2021). No estado do Amazonas, no norte do Brasil, a VOC Gamma também permaneceu predominante entre janeiro e maio de 2021, representando mais de 99% dos 752 sequenciados em um estudo (NAVECA et al., 2021b). Assim como observado nesses e outros estudos, a predominância da VOC Gamma do SARS-CoV-2 no país se refletiu no cenário de Botucatu-SP nos meses de abril e maio de 2021, período em que as amostras do presente estudo foram coletadas.

Além disso, é importante ressaltar que o primeiro semestre de 2021 no Brasil, especialmente após março, o número de casos e mortes aumentaram dramaticamente. Nessa época, em abril, o Brasil bateu o recorde mundial de mortes por Covid-19 em 24 horas, totalizando 4.250 vidas perdidas em um único dia (GIOVANETTI et al., 2021b). Enquanto no município de Botucatu, no mesmo mês, também foi registrado o recorde de mortes em 2021, sendo 15 óbitos. No mês seguinte, em maio, foi registrado o recorde de número de novos casos, totalizando 1.951 casos positivos (PREFEITURA MUNICIPAL DE BOTUCATU, 2021b). Nosso estudo abrangeu a metade final do mês de abril e a metade inicial do mês de maio, onde Gamma prevaleceu em 98,4% das amostras sequenciadas. Provavelmente essa onda de contágio nesse período, também foi causada e até impulsionada pela predominância da VOC Gamma.

Estima-se que Gamma pode ser até 2 vezes mais transmissível que as linhagens anteriores (Coutinho et al., 2021), além disso, soros convalescentes para o tipo selvagem podem sofrer uma perda de até 6,7x na capacidade de neutralização de P.1 (GIDARI et al., 2021). Esses achados também podem explicar a predominância de Gamma em Botucatu e no país, na época do estudo. Sobre a eficácia de vacinas contra a VOC Gamma, a redução dos títulos de neutralização em soros vacinados pela BNT162b2 (Pfizer) varia de 2,3x a 6,7x (GARCIA-BELTRAN et al., 2021; LUSTIG et al., 2021). E a eficácia clínica, pelo mesmo imunizante, chegou a 89% contra casos sintomáticos e 95% contra hospitalizações e óbitos (NASREEN et al., 2021). Nesse contexto, a vacinação ainda pode ser considerada a melhor estratégia para o manejo da Covid-19, uma vez que oferece uma maior proteção do que uma imunidade natural, além de possuir menos riscos à saúde humana.

Durante o estudo da presente pesquisa, também foram registradas algumas amostras contendo a VOC Alpha ou B.1.1.7 nas SE 17 (n=3) e SE 18 (n=2). A variante Alpha do SARS-CoV-2 foi identificada pela primeira vez no Reino Unido em 20 de setembro de 2020 e se disseminou por todo o país e posteriormente pela Europa continental (VOLZ et al., 2021; GAYMARD et al., 2021). A variante Alpha pode ser até 90% mais transmissível do que suas cepas ancestrais (DAVIES et al., 2021). Essa variante possui um agrupamento característico de mutações na proteína S, incluindo 7 substituições e 2 deleções (DAVIES et al., 2021; OSTROV, 2021).

Enquanto no presente estudo foram registrados apenas 5 genomas da VOC Alpha, no restante do Brasil, essa variante também foi registrada em uma porcentagem relativamente baixa quando comparada com a frequência de Gamma. A Rede de Alerta Pandêmico das Variantes Emergentes do SARS-CoV-2 detectou a VOC Alpha em 6,2% das 210 amostras selecionadas aleatoriamente no estado de São Paulo, até março de 2021. Gamma representou 66,2% desses genomas (SLAVOV et al., 2021a).

Em um outro estudo posterior, dos 21.700 genomas sequenciados pela mesma rede do estado, apenas 1,9% das sequências correspondia a VOC Alpha (SLAVOV et al., 2021b). No Brasil, até onde se tem conhecimento, o município de Ilhabela-SP foi o único lugar onde predominou Alpha. Entre o início de 2021 e o mês de junho (SE 24), 55,1% das sequências foram de genomas da VOC Alpha seguido por Gamma, que representou 41,5% do total de genomas sequenciados (INSTITUTO BUTANTAN, 2021a). Em um outro estudo mais abrangente (amostras coletadas entre maio e outubro de 2021), mais de 75% das sequências encontradas em Ilhabela, eram da variante Alpha (VIALA et al., 2022).



Além das análises de predominância das VOCs do SARS-CoV-2 no município de Botucatu antes da vacinação em massa, também foi analisada a dispersão das sublinhagens de Gamma ou P.1. As sublinhagens de uma variante do SARS-CoV-2 apresentam, além das mutações características de sua linhagem ancestral, mutações adicionais no seu genoma (IKETANI et al., 2022; LIMA et al., 2022).

As sublinhagens que conseguimos detectar de Gamma (P.1) foram P.1.14 e P.1.7. Foi registrada apenas uma sequência de P.1.7. As linhagens P.1 e P.1.14 se destacaram através das SE analisadas. A linhagem P.1 foi significativamente mais predominante na SE 16 do que a P.1.14. Nas semanas subsequentes, tanto P.1 como P.1.14 codominaram, não havendo diferenças estatisticamente significativa nas frequências das duas.

A sublinhagem P.1.14 tem como sua sequência mais antiga registrada em 03 de novembro de 2020 e foi encontrada mais comumente no Brasil, Canadá, Argentina, Estados Unidos das Américas e Japão. O maior número de sequências registradas foi entre março e agosto de 2021 (cov-lineages.org). No estado de São Paulo até meados agosto de 2022, segundo o Relatório de Linhagem P.1.14 disponível no site outbreak.info, 6,5% das 63.031 sequências produzidas são de P.1.14.

Além disso, fizemos uma breve análise para mensurar a frequência das variantes do SARS-CoV-2 pelo estado de São Paulo no mesmo período de coleta das amostras do presente estudo. Por meio dos dados fornecidos pela rede Rede de Alerta Pandêmico das Variantes Emergentes do SARS-CoV-2 e disponíveis no GISAID, pudemos observar que P.1 e P.1.14 foram as mais frequentes no estado. Quando separamos o estado pelas regiões administrativas da Secretaria de Estado da Saúde de São Paulo – DRS, curiosamente, a variante P.1 praticamente tão frequente quanto a sublinhagem P.1.14 na DRS de Bauru, região onde se localiza Botucatu. É importante mencionar que o número amostral (111 sequências) desse período pode estar subestimado, uma vez que essa DRS inclui vários municípios e, na SE 18, por motivos desconhecidos, não foram produzidas sequências. De qualquer forma, aparentemente essas duas linhagens parecem ter conduzido a segunda onda da Covid-19 não só no município de Botucatu, mas também na região do município especialmente.

Em um outro estudo realizado no Amazonas foi demonstrado que as sublinhagens de Gamma P.1.4 e P.1.6 foram as variantes mais prevalentes no estado durante o final do primeiro semestre de 2021 até setembro do mesmo ano (NAVECA et al., 2021c). Os valores médios de  $C_T$  também declinaram significativamente à medida que essas sublinhagens predominavam. Por um outro lado, a incidência dessas sublinhagens não aumentou em indivíduos vacinados em comparação com pessoas não vacinadas (NAVECA et al., 2022d).

Mesmo assim, é importante ressaltar que o aumento da incidência das sublinhagens de Gamma em comparação com P.1 pode ser explicado devido ao número de pessoas imunizadas (naturalmente infectadas e vacinadas) e que essas variantes podem estar associadas a uma maior infecciosidade ou que podem possuir mutações de evasão de anticorpos.

Na Índia, por exemplo, as sublinhagens de B.1.617, Delta e Kappa predominaram em 78% das sequências depositadas no GISAID entre abril e maio de 2021 (VENKATRAJA et al., 2022). Assim como observado nesse estudo na Índia, porém com outra linhagem do SARS-CoV-2, a segunda onda da Covid-19 no município de Botucatu-SP provavelmente foi conduzida pelas variantes P.1 e P.1.14 no mesmo período. É importante destacar que os meses em que foi realizado o presente estudo, foram registrados os maiores números de infecções e óbitos na cidade, como dito anteriormente.

Quando se correlaciona a Covid-19 com o sexo biológico dos pacientes, a literatura sugere que homens são mais propensos a desenvolver a doença na forma grave do que as mulheres (HAITAO et al., 2020; VIVEIROS et al., 2021). Essa diferença provavelmente pode ser explicada por uma enorme gama de fatores, tais como, o comportamento ou estilo de vida, a prevalência de comorbidades, os fatores genéticos sexuais e hormonais e as diferenças existentes nas vias biológicas relacionadas à infecção pelo SARS-CoV-2 entre os sexos.

Em nosso estudo, foi possível observar que, a linhagem P.1 foi significativamente mais frequente do que P.1.14 tanto em homens quanto em mulheres, sendo a proporção de sequências ainda maior em mulheres (60,3%;  $p < 0,0001$ ). No estado do Amazonas também foi observado um aumento no percentual de casos em mulheres após o surgimento da variante P.1 em relação ao surto causado por outra linhagem ancestral. Esses achados podem sugerir uma mudança no perfil de patogenicidade e virulência da variante em relação às suas ancestrais (FREITAS et al., 2021). Por um outro lado, as variantes que surgiram depois, especificamente Delta e Ômicron, foram sutilmente mais elevadas em pessoas do gênero masculino em um estudo de coorte nos EUA (WANG et al., 2022). Assim como os fatores relacionados ao gênero (comportamento, biológicos e genéticos, estilo de vida, por exemplo), o perfil genético das variantes e suas mutações também podem ser fatores considerados importantes na distribuição do SARS-CoV-2 entre os gêneros biológicos.

Também foram feitas análises sobre a proporção das variantes distribuídas em diferentes faixas etárias. Para isso, dividimos nossas amostras em três grandes grupos de acordo com a faixa etária dos participantes. A variante P.1 foi significativamente mais frequente do que P.1.14 nas pessoas jovens (0-19 anos) e nos adultos (20-59 anos). No grupo

com pessoas idosas (60 anos ou mais), tanto P.1 como P.1.14 foram estatisticamente semelhantes nas frequências.

Em um outro estudo realizado com amostras coletadas durante a segunda onda no estado do Amazonas, quando foi descoberta a variante P.1, houve uma maior incidência e aumento na proporção de casos de Covid-19 entre as pessoas mais jovens e adultos quando comparada com a primeira onda causada por outra variante (FREITAS et al., 2021). Isso também reforça que a variante P.1 pode estar associada com uma mudança no perfil da infecção pelo SARS-CoV-2 nas diferentes faixas etárias. Por um outro lado, a vacinação das pessoas idosas já havia sido iniciada no Brasil na mesma época deste estudo. Isso pode ser considerado importante na determinação da distribuição das variantes.

Sobre a frequência das variantes do SARS-CoV-2 em participantes que conviviam com um ou mais de um tipo de comorbidade, não conseguimos encontrar diferenças significantes. Entretanto é importante destacar que pessoas que convivem com alguma comorbidade são mais propensas a desenvolver formas mais graves da Covid-19 (FANG et al., 2020; VAGEESH; YUAN, 2020; TSANKOV et al., 2021). Nosso estudo possui algumas limitações amostrais para fazer inferências sobre a incidência das variantes nos casos com comorbidades. Nossa metodologia de seleção das amostras ocorreu de maneira randomizada, levando em conta qualquer pessoa contaminada da nossa lista. Dessa forma, entendemos que mais trabalhos relacionando a incidência das variantes do SARS-CoV-2 em indivíduos com comorbidades devem ser feitos com os critérios de seleção voltados para essas pessoas.

Quanto a sintomatologia da doença, a variante P.1 foi significativamente mais frequente nos pacientes possuindo algum sintoma da Covid-19 do que a sublinhagem P.1.14. Esses dados podem ter algum viés, visto que os pacientes que se testaram, procuraram o sistema de saúde local para fazer o teste provavelmente por algum motivo. Dessa forma, os pacientes assintomáticos podem estar subnotificados. Além disso, é importante mencionar a importância das vacinas contra o SARS-CoV-2 e suas variantes, que têm demonstrado reduzir significativamente a sintomatologia da Covid-19 (BERNAL et al., 2021a; BERNAL et al., 2021b; POUWELS et al., 2021; TREGONING et al., 2021).

É de conhecimento que os sintomas possuem relação com a carga viral (SORAAS et al., 2022). Entretanto, não conseguimos encontrar associação estatisticamente significativa nos valores de  $C_T$  entre as sublinhagens de Gamma (P.1 e P.1.14). É importante destacar que no presente estudo, esses valores foram obtidos de amostras que vieram de região de naso ou orofaringe, sem levar em consideração o momento da infecção.

As diferenças dos valores de  $C_T$  já foram observadas entre as VOC do SARS-CoV-2 em outros estudos. Adamoski et al. (2022) observaram que Gamma, possui um valor de  $C_T$  maior que o tipo selvagem ancestral. A variante Ômicron tem valores de  $C_T$  maiores do que Delta, e Delta tem valor maior do que Gamma. Em contrapartida, um outro estudo demonstrou que, em amostras de saliva, Alpha e Delta têm valores de  $C_T$  substancialmente menores do que variantes mais ancestrais (KING et al., 2022).

A variante P.1 foi significativamente mais frequente do que P.1.14 nos casos leves de Covid-19. Além disso, não houve diferenças significativas entre as proporções de casos moderados, graves e óbitos por Covid-19 em relação as variantes P.1 e P.1.14. A maior frequência de P.1 em casos leves provavelmente ocorreu porque, de fato, a maioria dos casos de Covid-19 geralmente é leve, e também, o maior número de sequências encontradas neste estudo foram da linhagem P.1.

Por um outro lado, o surgimento da variante P.1 no Amazonas foi fortemente relacionada com a grande onda de infecções e casos graves do estado (NAVECA et al., 2021a; FREITAS et al., 2021). No município de Botucatu, enquanto predominou a variante Gamma, na mesma época deste estudo, foram registrados os maiores números de internações e óbitos. Provavelmente esses números foram impulsionados pela VOC Gamma, incluindo a variante P.1 e sua sublinhagem P.1.14.

Antes do dia em que ocorreu a vacinação da população do estudo em Botucatu, algumas pessoas já tinham sido imunizadas de acordo com Plano Nacional de Operacionalização da Vacinação contra a Covid-19. No nosso estudo, 48 pessoas haviam sido imunizadas com CoronaVac (28) ou Astrazeneca (20). A variante P.1 foi significativamente mais frequente nas pessoas que tomaram a CoronaVac do que a sublinhagem P.1.14. Diferenças estatísticas não foram observadas entre as pessoas que tomaram o imunizante da AstraZeneca e entre os que tomaram uma ou duas doses destes imunizantes. De qualquer forma, é importante destacar que os números de casos e óbitos reduziram mais de 63% e 60% respectivamente entre a 1ª e a 2ª dose da vacinação massiva em Botucatu. A segunda dose aconteceu em 8 de agosto de 2021, e até o final do ano, os casos positivos reduziram mais de 91% e o número de vidas perdidas reduziram mais de 87% (PREFEITURA MUNICIPAL DE BOTUCATU, 2021b).

Quanto as análises filogenéticas realizadas neste estudo, pudemos observar que a maioria das amostras se agruparam no grande clado 20J (Gamma). Curiosamente as sequências de P.1 ficaram mais dispersas pela árvore, enquanto todas as sequências de P.1.14 ficaram agrupadas no ramo que contém a substituição na ORF1a: S2947N, o que é esperado

encontrar em P.1.14 (<https://outbreak.info/situation-reports?pango=P.1.14>). A maioria das sequências de P.1.14 possuem as substituições em ORF3a: K67N e T223I e até a data 12/08/2021 foram registradas 193 sequências de P.1.14 com esse genótipo, segundo o CovSpectrum ([cov-spectrum.org](https://cov-spectrum.org)). Investigando ainda mais esse ramo, conseguimos encontrar um pequeno agrupamento de 9 amostras de P.1.14 contendo a substituição em ORF1a: M584V e A3620V. Até o momento (12/08/2021) não foi registrada nenhuma sequência desta sublinhagem de Gamma no CovSpectrum, entretanto é importante considerar que encontramos uma quantidade relativamente baixa de amostras contendo essas mutações.

## 5 CONSIDERAÇÕES FINAIS

Nossos achados sugerem que a Variante de Preocupação (VOC) do SARS-CoV-2 Gamma (incluindo a linhagem P.1 e a sublinhagem P.1.14) foi predominante antes da vacinação massiva que ocorreu em Botucatu-SP com o imunizante ChAdOx1-nCov-19 - AZD1222 (AstraZeneca), especificamente no período entre 18 de abril e 15 de maio de 2021 (Semanas Epidemiológicas 16, 17, 18 e 19). A VOC Alpha e a Variante de Interesse (VOI) Zeta foram registradas, porém em baixa escala.

Ao comparar as frequências das linhagens P.1 e P.1.14, no que diz respeito a distribuição epidemiológica, a linhagem P.1 foi significativamente mais frequente tanto em homens quanto em mulheres. Essa mesma variante também foi a mais frequente entre jovens e adultos (0-59 anos), enquanto entre os idosos, P.1 e P.1.14 foram estatisticamente semelhantes nas frequências. Não foram observadas diferenças estatísticas nas frequências de P.1 e P.1.14 entre as pessoas que convivem com uma ou mais comorbidades.

A variante P.1 foi mais frequente do que P.1.14 nas pessoas com algum tipo de sintoma da Covid-19. Além disso, não houve diferença estatística nos valores de  $C_T$  entre as linhagens P.1 e P.1.14. Nenhuma dessas variantes foram associadas com os quadros clínicos (leve, moderado e grave) da Covid-19 ou óbitos. A variante P.1 foi significativamente mais prevalente do que a P.1.14 nos 28 participantes que tomaram o imunizante CoronaVac, enquanto nenhuma diferença foi observada entre os imunizados com a vacina da Astrazeneca (20 pessoas).

Quanto a filogenia das sequências encontradas antes da vacinação massiva em Botucatu, as sequências de P.1 ficaram mais dispersas pelo grande clado de Gamma enquanto P.1.14 ficou agrupada no clado contendo a substituição em ORF1a: S2947N. Ainda dentro deste clado, encontramos 9 sequências contendo as substituições em ORF3a: K67N e T223I e até o momento da escrita deste manuscrito, nenhuma sequência de P.1.14 com essas substituições foram encontradas na plataforma do CovSpectrum.

Nesse sentido, a alta predominância da VOC Gamma, incluindo suas linhagens P.1 juntamente com P.1.14 em um cenário pré vacinação em massa pode nos fornecer insights sobre a evolução e epidemiologia molecular do SARS-CoV-2 e de suas VOCs emergentes. Dessa forma ressaltamos a importância da vigilância genômica do SARS-CoV-2 que pode ajudar a subsidiar as tomadas de decisões dos setores públicos no manejo da COVID-19.

## REFERÊNCIAS

- ADAMOSKI, D. et al. "SARS-CoV-2 Delta and Omicron Variants Surge in Curitiba, Southern Brazil, and Its Impact on Overall COVID-19 Lethality." **Viruses** v. 14, n. 4, p. 809, 2022. doi: 10.3390/v14040809.
- ALEEM A; AKBAR SAMAD, AB, SLENKER, AK. "Emerging Variants of SARS-CoV-2 And Novel Therapeutics Against Coronavirus (COVID-19)." 2021. In: STATPEARLS [Internet]. Treasure Island (FL): StatPearls Publishing; 2022, Jan. May 12, 2022. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK570580/>.
- ANICHINI, G. et al. "Neutralizing antibody response of vaccinees to SARS-CoV-2 variants." **Vaccines** v. 9, n. 5, 517, 2021.
- BATES, T. et al. "Neutralization of SARS-CoV-2 variants by convalescent and BNT162b2 vaccinated serum." **Nature communications** v. 12, n. 1, p. 1-7, 2021. doi: 10.1038/s41467-021-25479-6.
- BERNAL, J. et al. "Effectiveness of Covid-19 vaccines against the B. 1.617. 2 (Delta) variant." **New England Journal of Medicine**, v. 385, n. 7, p. 585-594, 2021b. doi: 10.1056/NEJMoa2108891.
- BERNAL, J. et al. "Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study." **BMJ** v. 373, n. 1088, 2021a. doi: 10.1136/bmj.n1088.
- BOEHM, E. et al. "Novel SARS-CoV-2 variants: the pandemics within the pandemic." **Clinical Microbiology and Infection** v. 27, n. 8, p. 1109-1117, 2021. doi: 10.1016/j.cmi.2021.05.022.
- CALLAWAY, E.; MALLAPATY, S. Novavax offers first evidence that COVID vaccines protect people against variants. **Nature**. v. 590, n. 7844, p. 17, 2021. doi: 10.1038/d41586-021-00268-9.
- CAMPBELL, F. et al. "Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021." **Eurosurveillance** v. 26, n. 24, :2100509, 2021. doi: 10.2807/1560-7917.ES.2021.26.24.2100509.

CEVIK, M. et al. "Virology, transmission, and pathogenesis of SARS-CoV-2." *BMJ* v. 371, m3862, 2020. doi: 10.1136/bmj.m3862.

CHAKRABORTY, C.; BHATTACHARYA, M.; SHARMA, A. "Present variants of concern and variants of interest of severe acute respiratory syndrome coronavirus 2: their significant mutations in S-glycoprotein, infectivity, re-infectivity, immune escape and vaccines activity." **Reviews in Medical Virology** v. 32, n. 2, e2270, p. 1-14, , 2021. <https://doi.org/10.1002/rmv.2270>.

COLLIER, D. et al. "Sensitivity of SARS-CoV-2 B. 1.1. 7 to mRNA vaccine-elicited antibodies." *Nature* v. 593, p. 136–141, 2021. <https://doi.org/10.1038/s41586-021-03412-7>.

COMBE, M.; SANJUAN, R. "Variation in RNA virus mutation rates across host cells." **PLoS pathogens** v. 10, n. 1, e1003855, 2014. doi: 10.1371/journal.ppat.1003855.

COUTINHO, R. et al. "Model-based estimation of transmissibility and reinfection of SARS-CoV-2 P. 1 variant." **Communications Medicine** v. 1, 48, p.1-8, 2021. <https://doi.org/10.1038/s43856-021-00048-6>.

DAVIES, N. et al. "Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1. 7 in England." *Science* v. 372, n. 6538, eabg3055, 2021. doi: 10.1126/science.abg3055.

DEJNIRATTISAI, W. et al. "Antibody evasion by the Brazilian P. 1 strain of SARS-CoV-2." **BioRxiv**, March 15, 2021. Doi: <https://doi.org/10.1101/2021.03.12.435194>.

DEMOLINER, M. et al. "Predominance of SARS-CoV-2 P. 1 (Gamma) lineage inducing the recent COVID-19 wave in southern Brazil and the finding of an additional S: D614A mutation." **Infection, Genetics and Evolution** v. 96, 105134, 2021. doi: 10.1016/j.meegid.2021.105134.

DUFFY, S. "Why are RNA virus mutation rates so damn high?." **PLoS biology**, v. 16, n. 8, e3000003, 2018. doi: 10.1371/journal.pbio.3000003.

EMBORG, H-D. et al. "Vaccine effectiveness of the BNT162b2 mRNA COVID-19 vaccine against RT-PCR confirmed SARS-CoV-2 infections, hospitalisations and mortality in prioritised risk groups." **MedRxiv**, June 2, 2021. doi: <https://doi.org/10.1101/2021.05.27.21257583>.



FANG, X. et al. "Epidemiological, comorbidity factors with severity and prognosis of COVID-19: a systematic review and meta-analysis." **AGING** (albany NY) v. 12, n. 13, 12493–12503, 2020. doi: 10.18632/aging.103579.

FREITAS, A. et al. "The emergence of novel SARS-CoV-2 variant P. 1 in Amazonas (Brazil) was temporally associated with a change in the age and sex profile of COVID-19 mortality: A population based ecological study." **The Lancet Regional Health-Americas**, v. 1, Sept, 1, 2021. doi: 10.1016/j.lana.2021.100021.

FROME, E.; KUTNER, M.; BEAUCHAMP, J. "Regression analysis of Poisson-distributed data." **Journal of the American Statistical Association** v. 68, n. 344, p. 935-940, 1973.

GARCIA-BELTRAN; WILFREDO F., et al. "Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity." **Cell**, v. 184, n. 9, p. 2372-2383, e9, Mar., 2021. doi: 10.1016/j.cell.2021.03.013.

GAYMARD, A. et al. "Early assessment of diffusion and possible expansion of SARS-CoV-2 Lineage 20I/501Y. V1 (B. 1.1. 7, variant of concern 202012/01) in France, January to March 2021." **Eurosurveillance** v. 26, n. 9, 2100133, 2021. doi: 10.2807/1560-7917.ES.2021.26.9.2100133.

GIDARI, A. et al. "Cross-neutralization of SARS-CoV-2 B. 1.1. 7 and P. 1 variants in vaccinated, convalescent and P. 1 infected." **Journal of Infection** v. 83, n. 4, p.467-472, 2021. doi: 10.1016/j.jinf.2021.07.019.

GIOVANETTI, M. et al. Replacement of the Gamma by the Delta variant in Brazil: impact of lineage displacement on the ongoing pandemic. **Virus evolution**, v. 8, n. 1, p. veac024, 2022. doi: 10.1093/ve/veac024

GIOVANETTI, M. et al. "Genomic epidemiology reveals how restriction measures shaped the SARS-CoV-2 epidemic in Brazil." **MedRxiv**, Oct. 14, 2021b. doi: <https://doi.org/10.1101/2021.10.07.21264644>

GORBALENYA, AE. et al. The species Severe acute respiratory syndrome related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. **Nat Microbiol** v. 5, n. 4, p. 536–544, 2020. Doi: <https://dx.doi.org/10.1038/s41564-020-0695-z>.

HAAS, E. et al. "Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data." **The Lancet** v. 397, n. 10287, p. 1819-1829, May, 5, 2021. doi: [https://doi.org/10.1016/S0140-6736\(21\)00947-8](https://doi.org/10.1016/S0140-6736(21)00947-8).

HADFIELD, J. et al. NextStrain: Real-time tracking of pathogen evolution. **Bioinformatics**. v. 34, n. 23, p. 4121-4123, 2018. doi: [10.1093/bioinformatics/bty407](https://doi.org/10.1093/bioinformatics/bty407).

HAITAO, T. et al. "COVID-19 and sex differences: mechanisms and biomarkers." **Mayo Clinic Proceedings**. v. 95, n. 10, p. 2189-2203, 2020. doi: [10.1016/j.mayocp.2020.07.024](https://doi.org/10.1016/j.mayocp.2020.07.024).

HARRISON, A.; LIN, T.; WANG, P. "Mechanisms of SARS-CoV-2 transmission and pathogenesis." **Trends in immunology**, v. 41, n. 12, p. 1100-1115, 2020. doi: [10.1016/j.it.2020.10.004](https://doi.org/10.1016/j.it.2020.10.004).

HOFFMANN, M. et al. "The Omicron variant is highly resistant against antibody-mediated neutralization: Implications for control of the COVID-19 pandemic." **Cell** v. 185, n. 3, p. 447-456, e11, 2022. doi: [10.1016/j.cell.2021.12.032](https://doi.org/10.1016/j.cell.2021.12.032).

HUANG, Z. et al. "A review of the safety and efficacy of current COVID-19 vaccines." **Frontiers of Medicine** v. 16, n. 1, p. 39-55, 2022. 1-17. doi: [10.1007/s11684-021-0893-y](https://doi.org/10.1007/s11684-021-0893-y).

IKETANI, S. et al. "Antibody evasion properties of SARS-CoV-2 Omicron sublineages." **Nature** v. 604, n. 7906, p. 553-556, 2022. doi: [10.1038/s41586-022-04594-4](https://doi.org/10.1038/s41586-022-04594-4).

INSTITUTO BUTANTAN. Boletim Epidemiológico da Rede de Alerta das Variantes do SARS-COV-2: Ilhabela-SP: Instituto Butantan, 2021. Disponível em: [https://butantan.gov.br/assets/arquivos/Covid/Boletim\\_epidemiologico/20\\_06\\_21\\_Relat%C3%B3rio\\_Sequenciamento\\_Ilhabela.pdf](https://butantan.gov.br/assets/arquivos/Covid/Boletim_epidemiologico/20_06_21_Relat%C3%B3rio_Sequenciamento_Ilhabela.pdf). Acessado em: 11/03/2022 às 16:25

JAIN, V.; YUAN, J-M. "Predictive symptoms and comorbidities for severe COVID-19 and intensive care unit admission: a systematic review and meta-analysis." **International Journal of Public Health** v. 65, n. 5, p. 533-546, 2020. doi: [10.1007/s00038-020-01390-7](https://doi.org/10.1007/s00038-020-01390-7).

KING, K. et al. "SARS-CoV-2 variants of concern Alpha and Delta show increased viral load in saliva." **Plos one** v. 17, n. 5, e0267750, 2022. doi: 10.1371/journal.pone.0267750. eCollection 2022.

KUZMINA, A. et al. "SARS-CoV-2 spike variants exhibit differential infectivity and neutralization resistance to convalescent or post-vaccination sera." **Cell host & microbe** v. 29, n. 4, p. 522-528.e2, 2021. doi: 10.1016/j.chom.2021.03.008.

LAMARCA, A. et al. "Genomic surveillance of SARS-CoV-2 tracks early interstate transmission of P. 1 lineage and diversification within P. 2 clade in Brazil." **PLoS Neglected Tropical Diseases** v. 15, n. 10, e0009835, 2021. Doi: <https://doi.org/10.1371/journal.pntd.0009835>.

LAZAREVIC, I. et al. "Immune Evasion of SARS-CoV-2 Emerging Variants: What Have We Learnt So Far?." **Viruses** v. 13, n. 7, 1192, 2021. doi: 10.3390/v13071192.

LESBON, J. et al. "Nucleocapsid (N) gene mutations of SARS-CoV-2 can affect real-time RT-PCR diagnostic and impact false-negative results." **Viruses** v. 13, n. 12, 2474, 2021. doi: 10.3390/v13122474.

LIMA, A. et al. "SARS-COV-2 Genomic Monitoring in the State of SÃO PAULO UNVEILS TWO EMERGING AY. 43 SUBLINEAGES." **Journal of Medical Virology**, v.94, n. 7, p. 3394-3398, 2022. doi: 10.1002/jmv.27674.

LU, R. et al. "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding." **The lancet** v. 395, n. 10224, p. 565-574, 2020. [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8).

LUSTIG, Y. et al. "Neutralising capacity against Delta (B. 1.617. 2) and other variants of concern following Comirnaty (BNT162b2, BioNTech/Pfizer) vaccination in health care workers, Israel." **Eurosurveillance** v. 26, n. 26, p. 2100557. doi: 10.2807/1560-7917.ES.2021.26.26.2100557, 2021.

MINISTÉRIO DA SAÚDE. Plano Nacional de Operacionalização da Vacinação contra a covid-19. 2022. Disponível em: <https://www.gov.br/saude/pt-br/coronavirus/publicacoes-tecnicas/guias-e-planos/plano-nacional-de-operacionalizacao-da-vacinacao-contracovid-19.pdf>. Acessado em: 14/06/2022.

NAQVI, A. et al. "Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach." **Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease** v. 1866, n. 10, 165878, 2020. doi: 10.1016/j.bbadis.2020.165878.

NASREEN, S. et al. "Effectiveness of COVID-19 vaccines against symptomatic SARS-CoV-2 infection and severe outcomes with variants of concern in Ontario." **Nature Microbiology** v. 7, n. 3, p.379-385, 2022. doi: 10.1038/s41564-021-01053-0.

NAVECA, F. et al. 'COVID-19 in Amazonas, Brazil, was Driven by the Persistence of Endemic Lineages and P.1 Emergence', **Nature Medicine**, v. 27, n. 7, p. 1230-1238, 2021a. doi: 10.1038/s41591-021-01378-7.

NAVECA (b), F. et al. "Emergence and spread of SARS-CoV-2 P. 1 (Gamma) lineage variants carrying Spike mutations  $\Delta$ 141-144, N679K or P681H during persistent viral circulation in Amazonas, Brazil." **Image** v. 420, n. 784, p. 75-77, 2021b.

NAVECA, F., et al. "The SARS-CoV-2 variant Delta displaced the variants Gamma and Gamma plus in Amazonas, Brazil." 2021c. <https://virological.org/t/the-sars-cov-2-variant-delta-displaced-the-variants-gamma-and-gamma-plus-in-amazonas-brazil/765>.

NAVECA, F. et al. "Spread of Gamma (P. 1) sub-lineages carrying Spike mutations close to the furin cleavage site and deletions in the N-terminal domain drives ongoing transmission of SARS-CoV-2 in Amazonas, Brazil." **Microbiology spectrum** v. 10, n. 1, e0236621, 2022d. doi: 10.1128/spectrum.02366-21.

NIID Japan. Brief report: New variant strain of sars-cov-2 identified in travelers from Brazil. coronavirus disease 4. <https://www.niid.go.jp/niid/en/2019-ncov-e/10108-covid19-33-en.html> 2021.

NOVAVAX. Novavax Announces Positive Data from Three Complementary Studies of COVID-19 Beta (B.1.351) Variant Strain Vaccine. Jun., 11, 2021. <https://ir.novavax.com/2021-06-11-Novavax-Announces-Positive-Data-from-Three-Complementary-Studies-of-COVID-19-Beta-B-1-351-Variant-Strain-Vaccine>.

OSTROV, D. "Structural consequences of variation in SARS-CoV-2 B. 1.1. 7." **Journal of cellular immunology** v. 3, n. 2, p. 103-108, 2021. doi: 10.33696/immunology.3.085.

PACHETTI, M. et al. "Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant." **Journal of Translational Medicine** v. 18, Article number: 179, 2020. doi: <https://doi.org/10.1186/s12967-020-02344-6>.

PAPANIKOLAOU, V. et al. "From Delta to Omicron: S1-RBD/S2 mutation/deletion equilibrium in SARS-CoV-2 defined variants." **Gene** (2022): 146134.

PLANAS, D. et al. "Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization." **Nature** 596.7871 (2021): 276-280.

PLITNICK, J. et al. "Whole-genome sequencing of SARS-CoV-2: assessment of the Ion Torrent AmpliSeq panel and comparison with the Illumina MiSeq ARTIC protocol." **Journal of clinical microbiology** 59.12 (2021): e00649-21.

POUWELS, K. et al. "Effect of Delta variant on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK." **Nature medicine** 27.12 (2021): 2127-2135.

PREFEITURA MUNICIPAL DE BOTUCATU (a). 2021. Botucatu terá vacinação em massa no dia 16 de maio. 2021a. Disponível em: <https://www.botucatu.sp.gov.br/portal/noticias/0/3/22909/botucatu-tera-vacinacao-em-massa-no-dia-16-de-maio/#:~:text=A%20Prefeitura%20de%20Botucatu%2C%20com,a%20Covid%2D19%20em%20Botucatu>. Acesso em: 28/02/2022 às 13:44.

PREFEITURA MUNICIPAL DE BOTUCATU (b). 2021. Coronavírus. Disponível em: <https://www.botucatu.sp.gov.br/portal/vigilancia-epidemiologica/3/coronavirus/>. Acessado em: 12/08/2022.

RAMAN, R.; PATEL, K.; RANJAN, K. COVID-19: unmasking emerging SARS-CoV-2 variants, vaccines and therapeutic strategies. **Biomolecules**, v. 11, n. 7, p. 993, 2021.

RASHEDI, R. et al. "COVID-19 vaccines mix-and-match: The concept, the efficacy and the doubts." **Journal of Medical Virology** 94.4 (2022): 1294-1299.

SATORRA, A.; BENTLER P. "A scaled difference chi-square test statistic for moment structure analysis." **Psychometrika** 66.4 (2001): 507-514.

SHEIKH, J. et al. Emerging genetic diversity among clinical isolates of SARS-CoV-2: Lessons for today. **Infection, genetics and evolution**, v. 84, p. 104330, 2020.

SINAN. Calendário Epidemiológico de 2021. 2021. Disponível em: <http://portalsinan.saude.gov.br/calendario-epidemiologico-2020/43-institucional/171-calendario-epidemiologico-2021>. Acessado em: 16/03/2021 às 15:02

SLAVOV (a) S., et al. "Genomic monitoring unveil the early detection of the SARS-CoV-2 B. 1.351 (beta) variant (20H/501Y. V2) in Brazil." **Journal of Medical Virology** 93.12 (2021): 6782-6787.

SLAVOV (b) S. et al. "Genomic monitoring of the SARS-CoV-2 B1. 1.7 (WHO VOC Alpha) in the Sao Paulo state, Brazil." **Virus Research** 308 (2022): 198643.

SORAAS, A. et al. "Breakthrough infections with the omicron and delta variants of SARS-CoV-2 result in similar re-activation of vaccine-induced immunity." **medRxiv** (2022).

TREGONING, J. et al. "Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape." **Nature Reviews Immunology** v. 21, n. 10, p. 626-636, 2021. doi: 10.1038/s41577-021-00592-1.

TSANKOV, B. et al. "Severe COVID-19 infection and pediatric comorbidities: a systematic review and meta-analysis." **International Journal of Infectious Diseases** v. 103, p. 246-256, 2021. <https://doi.org/10.1016/j.ijid.2020.11.163>.

VAN DORP, L. et al. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. **Infection, Genetics and Evolution**, v. 83, 104351, 2020. <https://doi.org/10.1016/j.meegid.2020.104351>.

VENKATRAJA, B.; SRILAKSHMINARAYANA, G.; KUMAR, B. "The Dominance of Severe Acute Respiratory Syndrome Coronavirus 2 B. 1.617 and Its Sublineages and Associations with Mortality during the COVID-19 Pandemic in India between 2020 and 2021." **The American Journal of Tropical Medicine and Hygiene** v. 106, n.1, 142-149, 2022. doi: 10.4269/ajtmh.21-0812.

VIALA, V. et al. "The divergent pattern of SARS-CoV-2 variant predominance and transmission dynamics in the Brazilian island of Ilhabela." **Viruses** v. 14, n.7, p. 1481, 2022. doi: 10.3390/v14071481.

VIGNUZZI, M.; WENDT, E.; ANDINO, R. "Engineering attenuated virus vaccines by controlling replication fidelity." **Nature medicine** v. 14, n. 2, p. 154-161, 2008. doi: 10.1038/nm1726.

VIGNUZZI, M. et al. "Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population." **Nature** v. 439, n. 7074, p. 344-348, 2006. doi: 10.1038/nature04388.

VIVEIROS, A. et al. "Sex differences in COVID-19: candidate pathways, genetics of ACE2, and sex hormones." **American Journal of Physiology-Heart and Circulatory Physiology** v. 320, n. 1, H296-H304, 2021. doi: 10.1152/ajpheart.00755.2020.

VOLZ, E. et al. "Assessing transmissibility of SARS-CoV-2 lineage B. 1.1. 7 in England." **Nature** v. 593, n. 7858, p. 266-269, 2021. doi: 10.1038/s41586-021-03470-x.

WALL, E. et al. "Neutralising antibody activity against SARS-CoV-2 VOCs B. 1.617. 2 and B. 1.351 by BNT162b2 vaccination." **The Lancet** v. 397, n.10292, p. 2331-2333, 2021. doi: 10.1016/S0140-6736(21)01290-3.

WANG, C. et al. The establishment of reference sequence for SARS-CoV- 2 and variation analysis. **Journal of Medical Virology**, v. 92, n. 6, p. 667-674, 2020. doi: 10.1002/jmv.25762.

WANG, L. et al. "COVID infection rates, clinical outcomes, and racial/ethnic and gender disparities before and after Omicron emerged in the US." **MedRxiv** Feb 22;2022.02.21.22271300. 2022. doi: 10.1101/2022.02.21.22271300.

WHO. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. March, 2020a. <https://www.who.int/publications/m/item/covid-19-weekly-epidemiological-update>. Acesso em 28/02/2020 às 12:48.

WHO. World Health Organization. Coronavirus disease (COVID-19) Situation Dashboard. 2022b. Disponível em: <https://covid19.who.int/>. Acesso em: 07/04/2022 às 15:27.

WHO. World Health Organization. Weekly epidemiological update - 25 February 2021c. 2021. <https://www.who.int/publications/m/item/covid-19-weekly-epidemiological-update>. Acesso em: 28/02/2022 às 12:42.

WINK, P. et al. "Genomic Surveillance of SARS-CoV-2 Lineages Indicates Early Circulation of P. 1 (Gamma) Variant of Concern in Southern Brazil." **Microbiology spectrum** v. 10, n.1, e01511-21, 2022. <https://dx.doi.org/10.1128/spectrum.01511-21>.

WUHAN MUNICIPAL HEALTH COMMISSION. Report of clustering pneumonia of unknown etiology in Wuhan City. Wuhan Municipal Health Commission, 2019. <http://wjw.wuhan.gov.cn/front/web/showDetail/2019123108989>. Opens in new tab. Acesso em 11/04/2020 às 05:00.

ZAKI, A. et al. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. **New England Journal of Medicine**, v. 367, n. 19, p. 1814-1820, 2012. doi: 10.1056/NEJMoa1211721.

ZHANG, Y-Z. HOLMES, Edward C. "A genomic perspective on the origin and emergence of SARS-CoV-2." **Cell** v. 181, n. 2, p. 223-227, 2020. doi: 10.1016/j.cell.2020.03.035.

ZHONG, N. et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. **The Lancet**, v. 362, n. 9393, p. 1353-1358, 2003. doi: 10.1016/s0140-6736(03)14630-2.

ZHU, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. **New England Journal of Medicine**, v. 382, n. 8, p. 727-733. 2020. doi: 10.1056/NEJMoa2001017.