

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE MICROBIOLOGIA, IMUNOLOGIA E PARASITOLOGIA
CURSO DE GRADUAÇÃO EM FARMÁCIA

Geovana Peccatti Daniel

**Vigilância Genômica de SARS-CoV-2 em Florianópolis, Santa Catarina, Durante
a Terceira Onda da Pandemia da COVID-19**

Florianópolis
2023

Geovana Peccatti Daniel

**Vigilância Genômica de SARS-CoV-2 em Florianópolis, Santa Catarina, Durante
a Terceira Onda da Pandemia da COVID-19**

Trabalho de Conclusão de Curso submetido ao curso de graduação em Farmácia do Centro de Ciências da Saúde da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Bacharel em Farmácia.

Orientador: Prof. Dr. Glauber Wagner
Coorientadora: MSc. Dayane Azevedo Padilha

Florianópolis

2023

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Daniel, Geovana Peccatti

Vigilância Genômica de SARS-CoV-2 em Florianópolis, Santa Catarina, Durante a Terceira Onda da Pandemia da COVID-19 / Geovana Peccatti Daniel ; orientador, Glauber Wagner, coorientador, Dayane Azevedo Padilha, 2023.

52 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro de Ciências da Saúde, Graduação em Farmácia, Florianópolis, 2023.

Inclui referências.

1. Farmácia. 2. COVID-19. 3. Vigilância Genômica. I. Wagner, Glauber . II. Padilha, Dayane Azevedo. III. Universidade Federal de Santa Catarina. Graduação em Farmácia. IV. Título.

Geovana Peccatti Daniel

Vigilância Genômica de SARS-CoV-2 em Florianópolis, Santa Catarina, Durante a Terceira Onda da Pandemia da COVID-19

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de Bacharel em Farmácia e aprovado em sua forma final pelo Curso de Farmácia.

Florianópolis, 30 de junho de 2023.



Coordenação do Curso

Banca examinadora

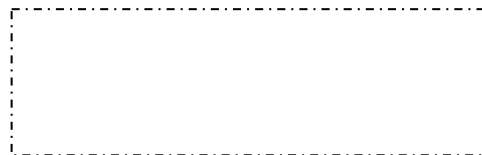


Prof. Dr. Glauber Wagner

Orientador



Dra. Doris Sobral Marques Souza
Universidade Federal de Santa Catarina



Profa. Dra. Maria Luiza Bazzo
Universidade Federal de Santa Catarina

Florianópolis, 2023

Este trabalho é dedicado a todos os que foram afetados ou perderam suas vidas em decorrência da COVID-19.

AGRADECIMENTOS

Agradeço a todos que de alguma forma contribuíram para que este projeto fosse concretizado. Primeiramente à minha família, minha mãe Cristina, minha irmã Diana, além de Diogo, Deise e Mércia, meu alicerce em todos os momentos. Sem o apoio de vocês nada seria possível. Aos meus colegas e amigos da graduação, agradeço por terem tornado minha experiência na faculdade tão especial.

Agradeço a todos os colegas do Laboratório de Bioinformática pelo acolhimento e parceria. Ao Eric e ao Vilmar, por terem compartilhado comigo o conhecimento e habilidades da bioinformática, sempre com muita paciência e disposição. Agradeço especialmente a Dayane, minha coorientadora, por todo o seu apoio e dedicação, que foi fundamental para o desenvolvimento deste projeto.

Ao professor Glauber, agradeço por ter me recebido no laboratório, me concedendo a oportunidade de conviver e aprender sobre esta área, bem como por confiar em mim uma parte desse projeto tão importante.

Minha gratidão se estende a todos aqueles que me proporcionaram oportunidades de aprendizado e experiências ao longo da graduação: colegas de estágio e de laboratório, professores e todos os outros que cruzaram meu caminho e se tornaram parte integrante desta trajetória.

Agradeço ao Laboratório de Protozoologia; ao Núcleo de Bioinformática; ao Laboratório de Biologia Molecular, Microbiologia e Sorologia; ao Laboratório de Virologia Aplicada e à empresa Mendelics Análise Genômica SA pela colaboração e por fornecerem estrutura e suporte para a realização deste trabalho.

Expresso minha gratidão também às contribuições do Projeto Sequenciamento em Rede Brasil - Rede Seqv; do Grupo Mulheres do Brasil; do Laboratório Central de Saúde Pública de Santa Catarina (LACEN-SC); da Diretoria de Vigilância Epidemiológica (DIVE); da Secretaria de Estado da Saúde de Santa Catarina (SES/SC), tanto na realização deste trabalho quanto pela dedicação no combate à pandemia de COVID-19.

Às instituições que forneceram subsídios para que este projeto fosse possível, como a Magazine Luíza, CAPES e FAPESC, meu muito obrigada.

RESUMO

O cenário da pandemia por SARS-CoV-2, vírus causador da COVID-19, foi marcado em 2021 pela troca de dominância da variante *Gamma* para *Delta*, seguido por uma diminuição do número de casos e de óbitos. A introdução da variante *Ômicron* no Brasil causou uma mudança no perfil de prevalência dos casos, aumentando expressivamente o número de notificações da doença. Neste contexto, técnicas de vigilância genômica se fizeram importantes para conhecer a genética do vírus e avaliar as variantes e linhagens, além de sua dinâmica de transmissão. Diante disto, foi avaliada a vigilância genômica do SARS-CoV-2 na região de Florianópolis, no estado de Santa Catarina, tendo em vista importância socioeconômica, turística e de alta concentração populacional desta localização. Foram analisados 338 genomas completos de SARS-CoV-2 obtidos de amostras coletadas entre novembro de 2021 e abril de 2022 nesta região, utilizando a tecnologia Oxford Nanopore (ONT). As variantes mais frequentes encontradas neste estudo foram da *Delta* (AY.101) e *Ômicron* (BA.1.1), e todos os genomas analisados tiveram cobertura de mais de 90%. Foi observado um padrão similar entre a distribuição das linhagens neste período quando comparado às encontradas em Santa Catarina e no Brasil. Os resultados deste trabalho contribuem para o conhecimento das variantes e linhagens de SARS-CoV-2 que estavam circulando em Florianópolis durante a terceira onda da COVID-19. Ainda, este estudo demonstra a importância do investimento em técnicas de vigilância genômica no controle da pandemia da COVID-19.

Palavras-chave: Vigilância genômica; COVID-19; *Ômicron*; Tecnologia Oxford Nanopore (ONT).

ABSTRACT

The scenario of the SARS-CoV-2 pandemic, the virus that causes COVID-19, was marked in 2021 by the dominance shift from the *Gamma* variant to the *Delta* variant, followed by a decrease in the number of cases and deaths. The introduction of the *Omicron* variant in Brazil caused a change in prevalence cases, significantly increasing the number of disease notifications. In this context, genomic surveillance techniques played an important role in understanding the genetics of the virus and assessing the variants and lineages, as well as their transmission dynamics. Therefore, genomic surveillance of SARS-CoV-2 was evaluated in the Florianópolis region, in the state of Santa Catarina, considering its socioeconomic importance, tourist appeal and high population density. A total of 338 complete SARS-CoV-2 genomes obtained from samples collected between November 2021 and April 2022 in this region were analyzed using Oxford Nanopore Technology (ONT). The most frequent variants found in this study were *Delta* variant (AY.101) and *Omicron* variant (BA.1.1), and all the analyzed genomes had coverage values above 90%. A similar pattern was observed in the distribution of lineages during this period when compared to those found in Santa Catarina and Brazil. The results of this study contribute to the understanding of the variants and lineages of SARS-CoV-2 that were circulating in the Florianópolis region during the third wave of COVID-19. This study demonstrates the importance of investing in genomic surveillance techniques in controlling the COVID-19 pandemic.

Keywords: Genomic surveillance; COVID-19; *Omicron*; Oxford Nanopore Technology (ONT).

LISTA DE FIGURAS

Figura 1- Representação da estrutura e genoma de SARS-CoV-2	17
Figura 2 - Esquema do Fluxo de trabalho do sequenciamento de genoma viral pela plataforma MinION Mk1C® (Oxford Nanopore Technologies, Oxford, Reino Unido).	22
Figura 3: Esquema de funcionamento da proteína do nanoporo.	27
Figura 4 - Amostragem total por período	36
Figura 5 - Frequência Relativa por Variantes por mês	37
Figura 6 - Frequência Relativa por Linhagens por mês	38
Figura 7 - Contagem da frequência de Variantes por mês	39
Figura 8 - Gráfico de escala multidimensional (MDS) da matriz de dissimilaridade de Jaccard	40
Figura 9 - Gráfico de escala multidimensional (MDS) da matriz de dissimilaridade de Jaccard	41

LISTA DE ABREVIATURAS E SIGLAS

CoV	Coronavírus
COVID-19	<i>Coronavirus disease 2019</i> - Doença de Coronavírus 2019
Cq	<i>Quantification cycle</i> - Ciclo de Quantificação
GISAID	<i>Global Initiative on Sharing All Influenza Data</i> - Iniciativa Global de Compartilhamento de Todos os Dados da Influenza
ICTV	<i>International Committee on Taxonomy of Viruses</i> - Comitê Internacional de Taxonomia de Vírus
IFN- β	Interferon beta
NGS	<i>Next generation Sequencing</i> - Sequenciamento de próxima geração
NSP	<i>Non-Structural Proteins</i> - Proteínas não estruturais
ONT	<i>Oxford Nanopore Technology</i> - Tecnologia Oxford Nanopore
ORF	<i>Open reading frame</i> - Quadro aberto de leitura
OMS	Organização Mundial da Saúde
RT-PCR	<i>Reverse transcription - polymerase chain reaction</i> - Transcrição reversa – reação em cadeia da polimerase
SARS	Síndrome Aguda Respiratória Severa
VOC	<i>Variant of Concern</i> - Variante de Preocupação
VOI	<i>Variant of Interest</i> - Variante de Interesse
VUM	<i>Variant under monitoring</i> - Variante em monitoramento

SUMÁRIO

1 INTRODUÇÃO	16
1.1 ORIGEM E ESTRUTURA DO VÍRUS	16
1.2. VIGILÂNCIA GENÔMICA	19
1.3. SEQUENCIAMENTO GENÉTICO	25
2 JUSTIFICATIVA	31
3 OBJETIVOS	32
3.1 OBJETIVO GERAL	32
3.2 OBJETIVOS ESPECÍFICOS	32
4 MATERIAIS E MÉTODOS	33
4.1 AMPLIFICAÇÃO E PREPARO DAS BIBLIOTECAS	33
4.1.1 Seleção das amostras	33
4.1.2 Síntese de cDNA	33
4.1.3 Preparo de Bibliotecas	33
4.2 ANÁLISE DE BIOINFORMÁTICA	34
4.3 ASPECTOS ÉTICOS	35
5 RESULTADOS	36
5.1 PERFIL DAS VARIANTES DE SARS-CoV-2 EM FLORIANÓPOLIS DURANTE O PERÍODO DE NOVEMBRO DE 2021 A ABRIL DE 2022	36
5.2 MUDANÇA DO PERFIL DE VARIANTES (DELTA PARA ÔMICRON)	37
6 DISCUSSÃO	42
7 CONCLUSÃO	48
REFERÊNCIAS	49

1 INTRODUÇÃO

1.1 ORIGEM E ESTRUTURA DO VÍRUS

SARS-CoV é classificado como um coronavírus (CoV) causador da Síndrome Aguda Respiratória Severa (SARS) que, em 2003, originou um surto emergente na China. Os coronavírus foram assim classificados porque possuem em sua estrutura externa proteínas de superfície que lembram o aspecto de uma coroa (corona, em latim) (ICTV, 2022).

Estes vírus possuem uma série de hospedeiros mamíferos, podendo ser transmitidos a humanos por mecanismos variados. Evidências apontam que SARS-CoV cruzou esta barreira intraespécie a partir do manuseio de espécies exóticas pela população em mercados de uma cidade da China (ICTV, 2022).

Os coronavírus fazem parte da família *Coronaviridae*, subfamília *Orthocoronavirinae* e que se dividem em quatro gêneros com diferentes afinidades patógeno-hospedeiro: Alphacoronavirus, Betacoronavirus, Gammacoronavirus e Deltacoronavirus. Enquanto os gêneros Gamma e Delta são exclusivamente zoonóticos, os Alpha e Betacoronavirus prevalecem entre os coronavírus responsáveis pelas doenças reportadas em humanos. Dos componentes destes dois últimos gêneros, apenas sete infectam os seres humanos, podendo causar doenças (CUI et al., 2019, CHEN et al., 2020).

Dos sete causadores de doença em humanos, quatro causam resfriados leves e com características que os tornam altamente transmissíveis, caracterizando-se por atingir somente a região nasal e da garganta — HCoV-NL63, HCoV-229E, HCoV-OC43 e HKU1. Ainda, dois são mais graves e infectam o trato respiratório baixo como os pulmões: o Coronavírus SARS (SARS-CoV) e o Coronavírus da síndrome respiratória do Oriente Médio (MERS-CoV) (FORNI et al, 2017).

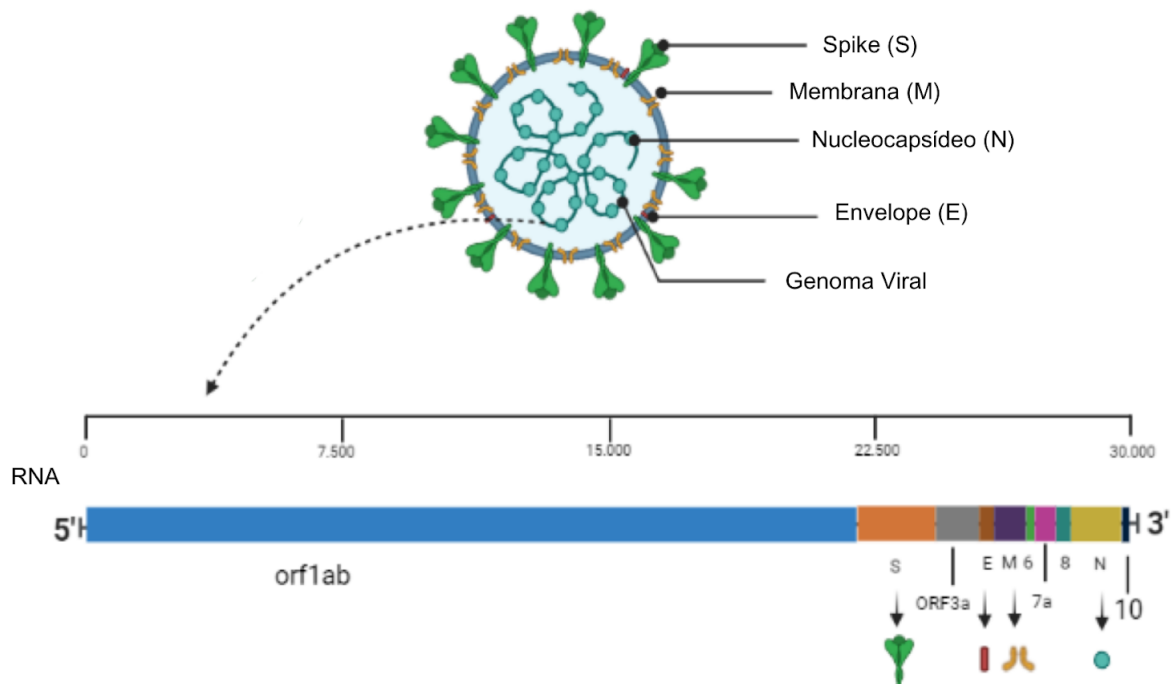
Por fim, mais um coronavírus foi identificado em Wuhan, na China, em dezembro de 2019, a partir de uma forma de pneumonia grave, sendo nomeado posteriormente como SARS-CoV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2*). É causador da “doença de coronavírus 2019”, a COVID-19 (*Coronavirus disease 2019*) (ZHU et al., 2020; WHO, 2020).

Uma das diferenças da família *Coronaviridae* em relação aos demais vírus de RNA (*ribonucleic acid*, em português ácido ribonucleico) encontra-se no seu genoma. Genoma é o conjunto de informações do material genético de um organismo. Estas informações são armazenadas na forma de ácidos nucleicos de cadeia simples ou dupla, dispostos em sequências lineares ou circulares, em uma cadeia de bases nitrogenadas (GIANI, et al., 2017).

Segundo o International Committee on Taxonomy of Viruses (ICTV), os *Coronaviridae* são os vírus de RNA com os maiores genomas identificados até o momento, com tamanho variando entre 29,8 kb a 29,9 kb, em comparação com o genoma do vírus Influenza, por exemplo, que é de 13,5 kb (GHEDIN, 2005; ICTV, 2022).

Os coronavírus são vírus envelopados com um genoma de RNA de sentido positivo de fita simples (ssRNA+). Seu genoma codifica 16 proteínas não estruturais (NSP), 4 proteínas estruturais e 6 proteínas acessórias (KHAILANY, et al., 2020).

Figura 1- Representação da estrutura e genoma de SARS-CoV-2



Adaptado de SCHLEMPER JUNIOR et al., 2022.

Os quadros de leitura aberta (ORFs) 1a e 1b codificam as poliproteínas que são posteriormente clivadas nas 16 NSPs. Estes são seguidas pelos genes que codificam as proteínas estruturais na seguinte ordem, conforme a nomenclatura: S,

E, M, e N (Figura 1). Estes genes codificam, respectivamente: a proteína *Spike* (S); a proteína Envelope (E); a proteína Membrana (M) e a proteína Nucleocapsídeo (N). A região que codifica as proteínas acessórias encontra-se entre S e E e entre M e N (BRANT et al., 2021).

Ambas as classes de proteínas citadas desempenham variadas funções essenciais para a manutenção da integridade e replicação viral. Entre as proteínas estruturais, a M é uma glicoproteína transmembranar que constitui o componente principal do envelope viral, além de ser a proteína estrutural mais abundante. Ela é crucial para a formação, liberação e empacotamento do material genético viral (ISLAM, et al., 2023).

A proteína E também é encontrada na bicamada lipídica do vírus, e tem a capacidade de formar poros que afetam a homeostase iônica das células hospedeiras, potencialmente contribuindo para a patogenicidade viral. Adicionalmente, a proteína E exibe propriedades de permeabilização de membrana celular (ISLAM, et al., 2023).

A proteína N é capaz de contribuir para a inibição da indução da citocina Interferon β (IFN- β) do hospedeiro pelo vírus. Ela também desempenha um papel essencial na expressão gênica de SARS-CoV-2 e de outros coronavírus. A proteína N configura-se como um alvo para diagnóstico e tratamento, pois, em comparação com outras proteínas alvo, possui uma sequência altamente conservada, é menos propensa a mutações e induz uma resposta imune robusta no hospedeiro (KANNAN; SUBBARAM; ALI; KANNAN, 2020).

A proteína *Spike* é dividida em duas subunidades: S1, responsável pela ligação do vírus ao receptor, e S2, responsável pela fusão das membranas do patógeno e das células hospedeiras. A proteína *Spike* é fundamental para a invasão das células do hospedeiro, pois liga-se à ACE2 (enzima conversora de angiotensina 2), promovendo a fusão da membrana e permitindo a replicação viral (ICTV, 2022).

Foi demonstrado que, no caso de SARS-CoV-2, a proteína *Spike* adquiriu mutações que aumentam sua afinidade com esse receptor. Um exemplo disso é a mutação D614G, identificada no início de 2020. Consiste na substituição do aminoácido aspartato por uma glicina, apresentando correlação positiva entre a carga viral e a transmissibilidade do vírus (VOLZ, et al., 2021).

As proteínas não estruturais ORF1a e ORF1b compreendem os primeiros dois terços do genoma viral. Desempenham papel fundamental na replicação e

persistência do vírus e codificam uma variedade de NSPs, tais como: NSP1, NSP2, NSP3, NSP12, NSP13 e NSP14. Estas proteínas são reconhecidas por sua interação com o hospedeiro através de sua capacidade de auxiliar o vírus a evadir o sistema imunológico. Inibem a ativação do IFN- β e induzem a ativação de vias inflamatórias que promovem a tempestade de citocinas, um fator agravante da doença (WU et al., 2020).

Todas estas estruturas estão sujeitas à mutabilidade, ou seja, alteração de sequência nucleotídica. Embora as mutações ocorram com frequência, não necessariamente resultam em mudanças nas características fenotípicas do vírus. No entanto, podem promover diversidade genética, o que pode afetar a interação viral com o hospedeiro, bem como dificultar o desenvolvimento de terapias e vacinas eficazes (KORBER, 2020). A alta taxa de replicação viral em hospedeiros com sistemas imunológicos deficientes exerce pressão seletiva, favorecendo mutações que conferem vantagens adaptativas para a sobrevivência e disseminação viral (GRUBAUGH, 2020).

Variantes que carregam a mutação D614G na proteína S, por exemplo, foram um sinal de alerta associado a maior transmissibilidade. Desde a identificação desta mutação, múltiplas variantes de SARS-CoV-2 foram descritas e suas mutações apresentaram diferentes fenótipos, com diferentes impactos na saúde pública (ALEEM; SAMAD; VAQAR, 2023).

A compreensão da característica de mutabilidade e diversidade genética do vírus SARS-CoV-2 é essencial para adaptar as estratégias de saúde pública diante das variantes emergentes (GRUBAUGH, 2020). Inicialmente, é necessário realizar o monitoramento das variantes, principalmente aquelas que adquirem vantagens adaptativas, representando um risco potencial de disseminação viral e agravamento da saúde populacional, podendo impactar diretamente as políticas de saúde pública.

1.2. VIGILÂNCIA GENÔMICA

Desde o primeiro caso reportado de COVID-19 (WU et al., 2020; ZHOU et al., 2020) até junho de 2023, foram confirmados mais de 760 milhões de casos e quase 7 milhões de mortes em todo o mundo. Do total, o continente americano soma quase 3 milhões de mortes, sendo que os Estados Unidos da América registraram o maior número, de 1,1 milhão de vítimas até o momento. Até junho de 2023 o Brasil

registrou mais de 702 mil óbitos, e relatou um total de 37,6 milhões de casos reportados (WHO, 2023).

A maioria dos casos do Brasil foi registrada na região sudeste, totalizando 14,9 milhões de casos, seguida pela região Sul, com 7,9 milhões de casos. Em Santa Catarina, até junho de 2023, foram registrados mais de 2 milhões de casos e mais de 22 mil óbitos. A capital do estado, Florianópolis, teve até o momento 214 mil casos confirmados e mais de 1,3 mil óbitos (BRASIL, 2022).

Enquanto a imunização, principalmente por meio de vacinas, é uma das soluções mais eficazes para restrição da propagação de epidemias virais, a vigilância epidemiológica é um método preventivo para o aumento do número de casos de doenças infecciosas. A vigilância tem importante papel para controle de doenças há mais de um século, através da coleta de dados e estudos de incidência e prevalência, além de ser uma estratégia de saúde pública amplamente utilizada e aceita como parte da proposta de mitigação de uma gama de doenças (CDC, 2022).

Uma das estratégias da vigilância epidemiológica é a vigilância genômica e, segundo a Organização Mundial da Saúde (OMS), esta última transforma a ação de saúde pública por fornecer uma compreensão mais aprofundada dos patógenos, podendo este conhecimento ser estendido para uma abordagem global, aplicada a todos os organismos potencialmente epidêmicos (WHO, 2020).

A vigilância genômica é realizada principalmente por técnicas moleculares de sequenciamento genético que permitem determinar a ordem das bases nucleotídicas que compõem o DNA ou RNA de patógenos, fornecendo informações detalhadas sobre suas características (CDC, 2022).

Através do entendimento sobre a evolução genética e a dispersão viral, é possível identificar eventos de importação e rotas locais de transmissão através da correlação com metadados (WHO, 2021). Ainda com o estudo de sua propagação, pode-se obter também discernimento sobre o envolvimento de outras espécies, como no caso dos surtos por MERS-CoV (HAAGMANS, 2014).

Durante a pandemia da COVID-19, a estratégia de vigilância genômica foi adotada no estado de Santa Catarina através de uma portaria de 2021, que instituiu a Rede de Vigilância Genômica do vírus SARS-CoV-2. Dentre os objetivos estavam o monitoramento da diversidade e evolução do vírus, além de procurar aprimorar a tomada de decisão frente a emergências públicas. Neste âmbito, a vigilância sanitária do estado e o Laboratório Central de Saúde Pública (LACEN-SC),

coordenaram a distribuição das amostras obtidas entre as diferentes instituições responsáveis por realizar o sequenciamento (SANTA CATARINA, 2021a).

No caso do vírus SARS-CoV-2, o sequenciamento pode ser realizado através de amostras de swab nasofaríngeo coletadas de pessoas cujo teste diagnóstico de RT-PCR (Transcrição reversa – reação em cadeia da polimerase) detectou a presença de material genético do vírus. Uma vez reunidas, as informações acerca do genoma são coletivamente utilizadas para investigar a propagação do vírus, com a finalidade de estimar a prevalência das variantes em uma população, desenvolver diagnósticos e vacinas, além de complementar outras fontes de dados e investigações epidemiológicas (CDC, 2022a).

O genoma de SARS-CoV-2 foi caracterizado pela primeira vez em Wuhan, na China, em fevereiro de 2020 por Zhu e colaboradores, e é utilizado globalmente como referência para a montagem das sequências obtidas (ZHU et al., 2020). Este processo inicia com a amplificação do material genético, sequenciamento e, por último, a análise de bioinformática das sequências obtidas, conforme demonstrado na Figura 2.

Atualmente, o *Global Initiative on Sharing All Influenza Data* (GISAID disponível em: <https://gisaid.org>) se constitui como o principal repositório de sequências do genoma de SARS-CoV-2 (ZHU et al., 2020). De janeiro de 2020 a junho de 2023 foram publicadas mais de 15,6 milhões de genomas completos na plataforma (GISAID, 2022).

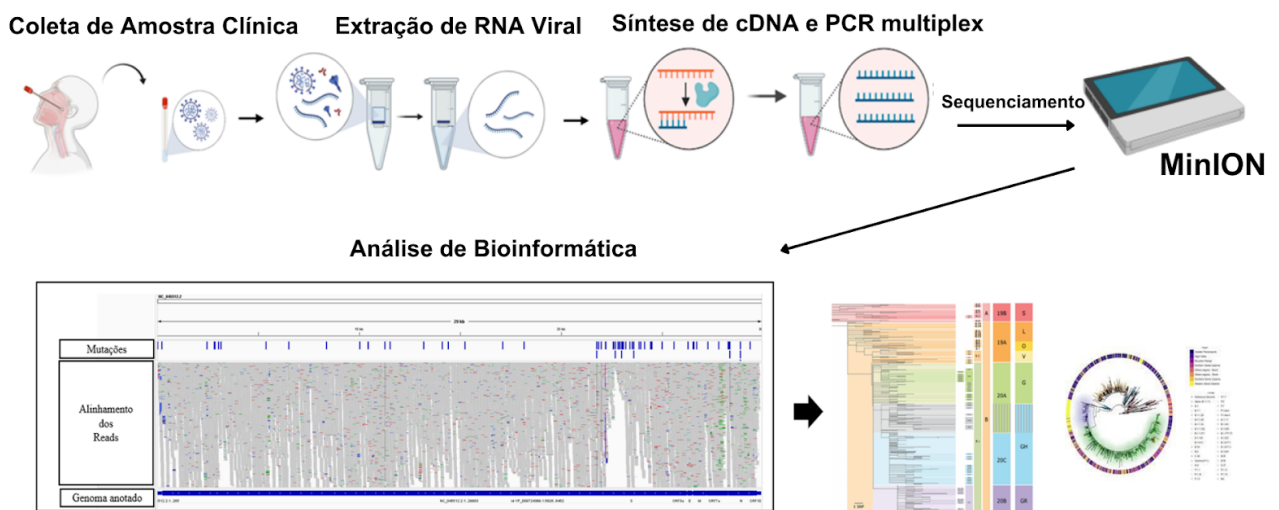
Os dados disponibilizados nesta plataforma são provenientes de 216 países, sendo que os Estados Unidos é o país com o maior número de sequências depositadas — mais de 4,7 milhões, seguido pelo Reino Unido com quase 3 milhões de sequências. O Brasil, até junho de 2023, compartilhou 227,5 mil sequências, representando 0,6% dos casos totais de COVID-19 confirmados (relação entre os casos reportados e o número de submissões de genoma no banco de dados). A mesma relação para Estados Unidos e Reino Unido é de 4,6% e 12%, respectivamente, evidenciando como o Brasil avançou pouco nesta área (GISAID, 2022).

Com base na compreensão sobre a evolução viral, é possível também avaliar o surgimento de variantes e linhagens circulantes. Isto permite identificar mutações e realizar estudos para avaliar a infectividade do patógeno, além de parâmetros como taxa de transmissão. Estes dados nos permitem prever com

antecedência o aumento de casos e possíveis surtos, permitindo estratégias de saúde pública mais direcionadas (WHO, 2021).

Os termos variante e linhagens estão ligados com a variação do genoma dentro de um mesmo organismo, e dizem respeito à intensidade da diferença. No nível mais fundamental ocorrem as mutações, que são alterações nucleotídicas, ou seja, de troca ou exclusão de pares de bases no momento de sua replicação. Nos vírus de RNA, como SARS-CoV-2, os mecanismos de reparo do material genético são escassos, o que aumenta consideravelmente a taxa de mutação quando comparado a vírus de DNA (PECK; LAURING, 2018).

Figura 2 - Esquema do Fluxo de trabalho do sequenciamento de genoma viral pela plataforma MinION Mk1C® (Oxford Nanopore Technologies, Oxford, Reino Unido).



Fonte: Adaptado de SCHLEMPER JUNIOR et al., 2022.

As mutações mais amplas podem ser consideradas variantes ou linhagens. Variantes possuem uma ou mais mutações específicas, que as classifica como um subtipo de um microrganismo, sem o considerar como uma nova cepa. Já as linhagens são atribuídas de acordo com a variante da qual descendem. São relacionadas a um ancestral comum e apresentam mutações similares, que não necessariamente estão associadas a características fenotípicas específicas (CDC, 2023).

Em junho de 2020, a OMS atribuiu nomenclaturas que identificavam as principais variantes do vírus SARS-CoV-2, objetivando priorizar o monitoramento de variantes específicas que representam um maior risco à saúde pública. Esta

classificação ainda persiste, e é feita considerando alterações significativas no funcionamento viral, além do fato de estarem circulando amplamente na população (WHO, 2023).

Dentre estas, foram classificadas as que possuem mutações que configuram Variantes de Interesse, do inglês *Variant of Interest* (VOI) e Variantes de Preocupação, do inglês *Variant of Concern* (VOC) (CDC, 2023). As primeiras são consideradas assim se, quando comparadas à variante original, tiverem alterações que a identificam como causadora de transmissão comunitária ou se tiver sido detectada em vários países. Exemplos de VOI são a *Epsilon*, da linhagem B.1.427; Zeta, da linhagem P.2; e *Kappa*, da linhagem B.1.617.1, todas designadas em 2021 (PAHO, 2021).

Já o grupo que se enquadra em *Variant of Concern* (VOC) também é classificado por uma avaliação comparativa, e é assim designado quando demonstra alterações associadas a um maior grau de significância para a saúde pública global. Dentre estes, podem ser destacados: (1) o aumento na transmissibilidade; (2) aumento na virulência; (3) mudança de aspectos clínicos da doença e (4) diminuição de eficácia de tratamento, diagnóstico e vacinas (PAHO, 2021).

Atualmente, não há mais nenhuma VOC em ascensão, somente VOI recombinantes descendentes da VOC *Ômicron*. Vale ressaltar que, devido ao cenário de variantes circulantes, em março de 2023, foi adicionada uma classificação independente somente de sublinhagens *Ômicron*, chamada de *Variants under monitoring* (VUMs) (WHO, 2023).

Posteriormente à primeira classificação da OMS foram contabilizadas inúmeras VOCs e VOIs, que moldaram as características da pandemia com relação à oscilação em números de casos, taxas de transmissibilidade e fatalidades, levando às características ondas da pandemia da COVID-19. Exemplo disso foi uma das mutações mencionadas anteriormente no genoma de SARS-CoV-2, que foi identificada na proteína *Spike*, a D614G, sendo esta relacionada à linhagem B.1, a qual aumentou a quantidade de carga viral e sua prevalência nos casos identificados (KORBER et al., 2020).

Estas mutações ocasionaram o surgimento de variantes em diferentes localidades do mundo e devido a diferentes pressões evolutivas. Segundo a OMS, as VOCs que circularam anteriormente durante a pandemia foram a *Alfa*, da

linhagem B.1.1.7; *Beta*, da linhagem B.1.351; *Gamma*, da linhagem P.1; *Delta*, da linhagem B.1.617.2 e *Ômicron*, da linhagem B.1.1.529 (WHO, 2022).

A VOC *Alfa* foi a primeira a ser descrita, a partir de uma amostra obtida no Reino Unido, em dezembro de 2020 (CHAND et al., 2020), e foi apontada como VOC pela OMS no mesmo período (CDC, 2023). As alterações mais significativas sofridas por ela foram encontradas na proteína *Spike*, que está envolvida na entrada do vírus nas células hospedeiras através da sua afinidade com o receptor ACE2.

Na variante *Alfa* foram detectadas sete mutações (RAMBAUT et al., 2020), as quais foram associadas ao aumento da capacidade do vírus de invadir as células hospedeiras, resultando em: aumento de carga viral, da frequência de detecção nos casos positivos e também ao aumento da mortalidade (DAVIES et al., 2021).

A variante *Beta* foi outra VOC inicialmente relatada na África do Sul, também em dezembro de 2020. Esta também tem duas mutações no domínio de ligação ao receptor, que aumentam a afinidade de ligação à enzima ACE2. Sua detecção ocorreu em meio a um cenário de aumento de casos, onde a associação ao aumento da transmissibilidade foi imediata. De outubro a dezembro de 2021, a predominância dela sobre o total de amostras sequenciadas saltou de 11% para 87% (KARIM, OLIVEIRA, 2021).

Em outubro de 2020, outra variante foi identificada e nomeada *Delta*. Foi reportada pela primeira vez na Índia (MLCOCHOVA et al., 2021), e em alguns meses já tinha se espalhado por 163 países (BHATTACHARYA, et al., 2023). As mutações observadas nesta variante afetaram tanto a interação com a enzima ACE2 quanto a eficácia dos anticorpos neutralizantes (BARNES, et al., 2020). Isso resultou em um aumento significativo na taxa de transmissão, que foi maior do que todas as variantes identificadas anteriormente (BHATTACHARYA, et al., 2023).

O sequenciamento genômico de amostras obtidas na região de Manaus, na região norte do Brasil, em novembro de 2020, revelou o surgimento de uma nova variante de SARS-CoV-2, a *Gamma*, da linhagem P.1. Esta acumulou 17 mutações de relevância fenotípica, incluindo na proteína *Spike*, e que foram associadas a uma maior ligação ao receptor humano ACE2 (FARIA et al., 2021).

As estimativas de transmissão da VOC *Gamma* foram, em média, 2 vezes maiores que as identificadas até o momento, além de causar infecções com cargas virais até 10 vezes mais altas quando comparadas às de outras linhagens (NAVECA et al., 2021). Dois meses após ser classificada como VOC, em janeiro de 2021

(WHO, 2023), foi predominante em mais de 90% das variantes identificadas neste período (GRAF et al., 2022).

Apesar de seu surgimento ter sido identificado no estado do Amazonas, sua transmissão também foi considerável em regiões muito diferentes geograficamente, como aconteceu em centros urbanos mais densos da região sudeste e sul. A densidade geográfica foi observada como um fator crítico na disseminação do vírus em diversas regiões do Brasil (GRAF et al., 2022).

No estado de Santa Catarina, foram comparadas as dinâmicas de transmissão da VOC *Gamma* em diferentes regiões do estado. Observou-se que a presença da variante *Gamma P.1-like-II* na região Oeste do estado estava associada ao aumento da propagação da doença e também à maior taxa de mortalidade (PADILHA et al., 2022).

A última variante a ser identificada foi a *Ômicron*, na África do Sul, em novembro de 2021 (VIANA, et al., 2022). A *Ômicron* acumula mais de 40 mutações, sendo 30 delas na proteína *Spike*, incluindo deleções que afetam a transmissibilidade e capacidade de evadir a resposta imune (WEI et al., 2021). O primeiro caso confirmado dessa variante no Brasil ocorreu no final de novembro de 2021, e sua detecção em Santa Catarina foi em dezembro do mesmo ano, e seu surgimento no estado ocorreu em Florianópolis (SANTA CATARINA, 2022b).

Após um ano do início da pandemia de COVID-19, as estratégias de vigilância genômica melhoraram significativamente na maioria dos países, incluindo um aumento no número de sequenciamentos e maior rapidez na submissão das informações ao GISAID. Na América do Sul o aumento das submissões de dados genômicos foi de 5,2 vezes (de 1 para 5 submissões a cada mil casos), enquanto o Brasil aumentou 6,5 vezes (de 0,9 para 5,6 a cada mil casos) (MAHANTA; SABERWAL; SHARMA, 2022).

Os dados apresentados ajudam a compreender o cenário através da identificação precisa das variantes de SARS-CoV-2, e são obtidos através das tecnologias de sequenciamento, como abordado anteriormente.

1.3. SEQUENCIAMENTO GENÉTICO

Quando Sanger lançou a primeira tecnologia de sequenciamento, em 1977, os custos por base eram elevados, tinham velocidade de processamento limitada e o

tamanho do genoma a ser sequenciado era limitado. Desde então, empreendimentos como o Projeto Genoma Humano impulsionaram avanços no desenvolvimento de métodos mais eficientes, e que permitem também aumento na velocidade e rendimento na geração dessas informações (GIANI, et al., 2017).

As tecnologias de sequenciamento que surgiram após os anos 2000 foram chamadas de sequenciamento de próxima geração ou NGS. Este tornou-se mais rápido e barato, sendo capaz de processar bilhões de fragmentos de DNA de forma simultânea. É dividido em segunda geração, que gera leituras curtas, e em terceira geração, que gera leituras longas (BARSKI, A. et al., 2021).

Entre as técnicas de sequenciamento de segunda geração encontram-se os sequenciadores de metodologia conhecida como sequenciamento por síntese, como a tecnologia Illumina, que utilizam uma técnica acurada, lendo fragmentos em torno de 150 pares de bases (YE et al., 2012). A terceira geração tem como representantes principais as metodologias de sequenciamento de molécula única (SMS) e o sequenciamento por nanoporos, que foi a metodologia utilizada neste trabalho.

O conceito de sequenciamento de nanoporos surgiu da junção de diferentes ideias na década de 1980, e foi viabilizado por uma série de avanços tecnológicos em suas proteínas motoras relacionadas. A alfa-hemolisina, uma proteína de canal de membrana de *Staphylococcus aureus* com um diâmetro interno de ~1,4 nm a ~2,4 nm, foi a primeira a mostrar que pode detectar diferentes sinais de corrente iônica pela passagem de bases nucleotídicas (WANG, et al., 2021).

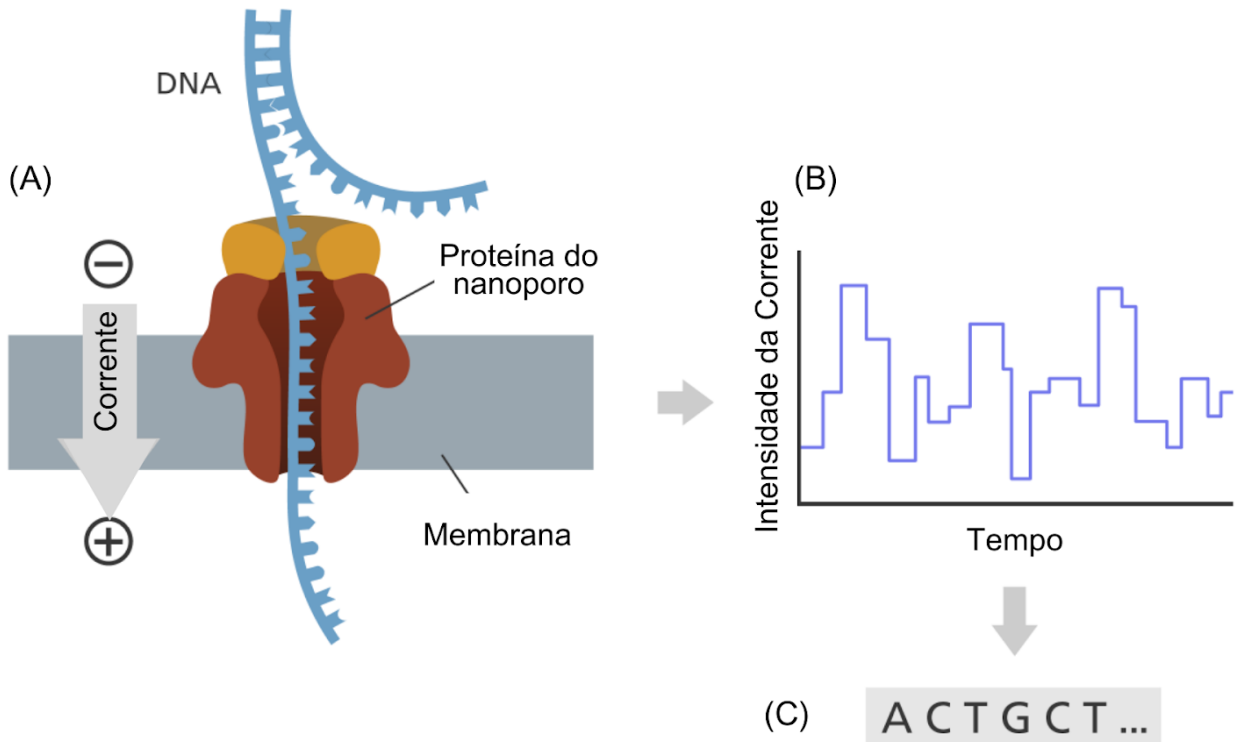
O nanoporo funciona como um biosensor, sendo a única passagem entre dois lados de uma membrana (Figura 3). Estes lados são o cis, composto por uma solução iônica e fragmentos do DNA a serem lidos; e o lado trans, carregado também ionicamente e para onde é direcionada a passagem da corrente elétrica. Uma vez que a condutividade iônica é sensível ao campo elétrico associado e à diferença de massa dos nucleotídeos, quando ocorre mudança na corrente elétrica através do nanoporo, é revelada a sequência que está sendo lida. Ademais, é possível sequenciar fragmentos de leitura com maior número de pares de bases, ou seja, leituras longas (DEAMER; AKESON; BRANTON, 2016).

Essa abordagem de sequenciamento foi utilizada pela empresa Oxford Nanopore Technologies (Oxford, Reino Unido), que lançou, entre outros, a tecnologia utilizada neste trabalho. A plataforma apresenta a vantagem de ser

portátil e mais acessível em termos de custo em comparação com outras tecnologias de sequenciamento de próxima geração (NGS). Ainda enfrenta desafios tendo em vista sua metodologia, porém já foram implementadas melhorias neste sentido, como a modificação na estrutura da proteína para retardar a translocação do DNA, que resulta em uma detecção mais precisa da molécula (DEAMER; AKESON; BRANTON, 2016).

Figura 3: Esquema de funcionamento da proteína do nanoporo.

(A) Estrutura representativa da tecnologia de sequenciamento por nanoporos, representando a proteína do nanoporo, a membrana que separa as regiões eletropositivas das eletronegativas, representação da passagem da corrente do polo negativo para o polo positivo e a passagem do material genético a ser sequenciado. (B) Esquema representativo de como o resultado do sequenciamento é gerado com base na intensidade da corrente que cada nucleotídeo gera ao passar pelo nanoporo com o passar do tempo. (C) Imagem representativa da sequência de bases geradas após o sequenciamento.



Adaptado de: yourgenome.org

A Tecnologia Oxford Nanopore (ONT) foi utilizada durante a epidemia do Ebola na Guiné, oeste do continente africano, em 2014 (QUICK et al., 2016), e no surto de Zika vírus que ocorreu no Brasil entre os anos de 2014 e 2015 (JESUS et al., 2019). Nesses casos, foi utilizada como uma alternativa para as tecnologias de sequenciamento tradicionais que requerem calibrações repetidas, disponibilidade de energia e de pessoal treinado (AREVALO, et al., 2022).

Dentre os parâmetros mais utilizados para determinar a qualidade do genoma, os mais frequentes são a cobertura e a profundidade. A cobertura é uma medida que se refere à proporção do genoma que foi sequenciada. Ela indica a extensão e o total do material genético que foi lido. Uma cobertura mais alta garante uma representação precisa de todo o genoma, no sentido de preencher possíveis lacunas (SIMS, et al., 2014).

A profundidade refere-se ao número de vezes que cada nucleotídeo é sequenciado durante o processo de sequenciamento. Uma maior profundidade implica múltiplas leituras de cada nucleotídeo individualmente, melhorando a confiabilidade e precisão dos resultados de sequenciamento (SIMS, et al., 2014). No caso da descoberta de variantes, esse indicador é fundamental porque garante que as variações encontradas refletem o material genético do vírus e não um erro de leitura.

Ambas as métricas desempenham papéis importantes no sequenciamento do genoma, uma vez que aumentam a precisão e a integridade dos dados obtidos. Uma maior profundidade e cobertura proporcionam maior confiança nos resultados do sequenciamento, permitindo identificar variações genéticas, e a compreensão do funcionamento dos organismos em um nível molecular (PETRACKOVA, et al., 2019).

Vários estudos abordaram a validação do sequenciamento por ONT em relação à sua acurácia. Dentre as abordagens, foram feitas comparações com a plataforma mais consolidada de *Whole-Genome Sequencing* (WGS) viral atualmente, a tecnologia Illumina. Estes estudos compararam critérios de qualidade e acurácia dos resultados, levantando questões sobre a cobertura baixa, taxas de erros e sensibilidade dessa tecnologia (TSHIABUILLA, et al., 2022). Porém, os resultados também demonstraram que é possível obter genomas consenso altamente precisos de SARS-CoV-2 (BULL, et al., 2020).

No Brasil a plataforma de sequenciamento MinION foi utilizada no surto de Zika vírus que ocorreu no Brasil entre os anos de 2014 e 2015, através do projeto ZiBRA (Zika in Brazil Real-time Analysis), como mencionado anteriormente. Este baseou-se em um exemplo anterior de um laboratório de vigilância genômica móvel durante o surto de Ebola na Guiné em 2014 (QUICK, et al., 2016), que também enfrentava um desafio de falta de dispositivos de sequenciamento para pesquisa de

campo. Durante este projeto foram gerados 54 genomas do vírus Zika, estabelecendo relações sobre o estabelecimento deste nas Américas e detectando inclusive a infecção mais antiga do país (JESUS et al., 2019).

O projeto ZiBRA também se constituiu por um esforço colaborativo de vigilância genômica, e o mesmo grupo de pesquisa foi responsável por sequenciar o primeiro genoma de SARS-CoV-2 brasileiro em fevereiro de 2020. O sequenciamento foi feito utilizando a tecnologia ONT e o mesmo fluxo de trabalho utilizado neste estudo. Através da análise filogenética foi possível avaliar a dinâmica de introdução de SARS-CoV-2 no Brasil, além das duas primeiras transmissões locais reportadas (JESUS, et al., 2020).

A tecnologia ONT ainda é estudada quanto à sua validação e acurácia em relação a tecnologias como Illumina, que são mais consolidadas. Entre os estudos que relacionam o uso dessa tecnologia com o sequenciamento do genoma de SARS-CoV-2, foi possível obter a montagem de genomas de consenso utilizando essa tecnologia. O estudo de Bull e colaboradores (2020) demonstrou uma determinação precisa de sequências de consenso com dados de ONT. Além disso, relatou adequação para as análises filogenéticas padrão e discute que embora a taxa de erro seja alta, essa tecnologia pode ser capaz de detectar variações estruturais que afetem funções gênicas do vírus (BULL, et al., 2020).

Em testes de comparação entre as duas tecnologias, foi demonstrado que a cobertura média entre os genomas obtidos teve uma diferença notável. Tshiabuila e colaboradores (2022) obtiveram em seu estudo uma média de cobertura de 94,3% e 72,9% para o sequenciamento pela plataforma Illumina e ONT, respectivamente. Foi relatado também que, enquanto 89% das amostras obtidas por Illumina passaram pelo controle de qualidade para submissão no banco de dados do GISAID, somente 27,9% das sequências obtidas pela ONT tiveram o mesmo resultado (TSHIABUILA et al., 2022).

Foi notada também uma maior cobertura genômica quando associada aos valores de *Quantification cycle* (Cq). Esse valor refere-se ao número de ciclos necessários para amplificar o RNA a um nível que seja detectável na técnica de PCR. Portanto, é inversamente proporcional à quantidade de material genético viral presente na amostra. Os achados de Tshiabuila e colaboradores (2022) demonstram que a plataforma Illumina não teve variações nos valores de cobertura genômica, enquanto o ONT teve resultado afetado por conta da diferença neste valor.

O fluxo de trabalho utilizado por Freed e colaboradores (2020), demonstrou alcançar resultados com uma cobertura de leitura de 99,9% do genoma e com profundidade de 50 vezes, avaliando amostras com valores de Cq entre 20 e 31, utilizando a tecnologia ONT. Isso foi alcançado com redução no tempo de preparo de biblioteca e no tempo de corrida de sequenciamento, além de menor custo atrelado (FREED et al. 2020).

A análise de bioinformática também foi padronizada e foi demonstrado que atinge resultados acurados, podendo ser realizada rapidamente. Algumas ferramentas utilizadas são protocolos da *ARTIC Network*, um projeto que fornece ferramentas para epidemiologia genômica de surtos e vírus (ARTIC NETWORK, 2023). Este inclui os primers utilizados para a amplificação do genoma de SARS-CoV-2, que é baseado no genoma de referência obtido em Wuhan (ZHU et al., 2020), além de programas de visualização em tempo real do sequenciamento, o RAMPART (*Read Assignment, Mapping, and Phylogenetic Analysis in Real Time*) (AREVALO, et al., 2022).

A atribuição de linhagens é feita utilizando o Pangolin (*Phylogenetic Assignment of Named Global Outbreak Lineages*), uma ferramenta proposta por O'Toole e colaboradores (2021), baseada no sistema de nomenclatura Pango (RAMBAUT et al. 2020). Esta define agrupamentos filogenéticos através de informações epidemiológicas relevantes, como a propagação de uma linhagem em uma área geográfica diferente, além de designar também a nomenclatura das VOCs (O'TOOLE et al., 2021).

A utilização deste conjunto de tecnologias configura-se como uma alternativa para monitorar o comportamento das variantes durante um surto epidêmico, buscando dados relevantes para o entendimento de aspectos de transmissão e patogenicidade de SARS-CoV-2. Com isto, a busca por alternativas que tendem a diminuir custos e agilizar o tempo no sequenciamento de patógenos infecto-contagioso torna-se importante.

2 JUSTIFICATIVA

A circulação de SARS-CoV-2 representa um desafio substancial para a sociedade, visto ser o vírus causador da pandemia que levou à morte de milhões de pessoas entre 2020 e 2023. Por este motivo, faz-se necessário compreender a dinâmica de comportamento das variantes na capital do estado de Santa Catarina e em cidades como Florianópolis. Isto devido a sua importância socioeconômica, turística e de alta concentração populacional. Este último fator é crítico na disseminação do vírus em diversas regiões do Brasil e neste estado, estando relacionado a taxas elevadas de transmissões (GRAF et al., 2022; PADILHA et al., 2022). Com isto, foi necessário o investimento em técnicas de sequenciamento, que objetivaram a vigilância genômica deste vírus a fim de acompanhar sua evolução, adaptação ambiental e impactos clínicos. Este estudo objetivou o uso de sequenciamento pela tecnologia ONT, conhecida técnica otimizada para a produção de dados de forma mais ágil e com menor custo. Quando comparado a outras tecnologias NGS, os dados gerados através da ONT também podem permitir a monitorização do comportamento do vírus e podem constituir como uma estratégia de vigilância genômica.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Avaliar as variantes e linhagens de SARS-CoV-2 em amostras obtidas em Florianópolis no período de novembro de 2021 a abril de 2022, a partir do sequenciamento utilizando a tecnologia Nanopore.

3.2 OBJETIVOS ESPECÍFICOS

- Sequenciar amostras biológicas humanas obtidas em Florianópolis, no período de novembro de 2021 a abril de 2022;
- Identificar as variantes e linhagens circulantes no período do estudo;
- Avaliar, comparativamente, as linhagens e mutações observadas neste período com amostras de outras regiões do Estado e do Brasil.

4 MATERIAIS E MÉTODOS

4.1 AMPLIFICAÇÃO E PREPARO DAS BIBLIOTECAS

4.1.1 Seleção das amostras

Para este estudo, foram sequenciadas 338 amostras obtidas a partir de swabs nasofaríngeos. Nestes, a detecção do genoma de SARS-CoV-2 por RT-PCR, apresentou *Quantification cycle* (Cq) em até 25. O material genético destas amostras foi cedido pelo Laboratório de Biologia Molecular, Microbiologia e Sorologia, do Centro de Ciências da Saúde/UFSC (LBMMS/CCS/UFSC). Foi proveniente de coletas realizadas entre o período de novembro de 2021 a abril de 2022, compreendendo 39 semanas epidemiológicas (da 31 de 2021 até 17 de 2022).

4.1.2 Síntese de cDNA

Para a etapa do sequenciamento, as amostras foram processadas no Laboratório de Protozoologia da UFSC. Este processo inicia com a transcrição reversa do RNA em cDNA, que é feita utilizando o reagente LunaScript® RT SuperMix (New England Biolabs, Massachusetts, USA).

Seguindo uma abordagem desenvolvida pela ARTIC Network, a PCR foi realizada pelo protocolo de sequenciamento nCoV-2019 *Midnight RT PCR Expansion (EXP-MRT 001)*, que inclui o painel de primers V3. Para cada amostra foram feitos dois *pools* com iniciadores diferentes, denominados A e B, sendo estes amplificados separadamente. Após a amplificação foi realizada a quantificação do material genético utilizando o *Qubit Fluorometric Quantification* (Thermo Fisher Scientific Inc, Massachusetts, USA) e os amplicons seguiram no processo de preparo das bibliotecas para o sequenciamento NGS e purificação.

4.1.3 Preparo de Bibliotecas

O preparo da biblioteca de sequenciamento foi feito em três etapas principais: ligações do código de barras - que identifica a amostra- purificação do material genético por *beads* e ligação do adaptador de sequenciamento. Para tanto, se utilizou o *Rapid Barcoding Kit 96* (SQK-RBK110.96) e o *Ligation Sequencing Kit*, LSK-109 (Oxford Nanopore Technologies, Oxford, Reino Unido), fornecidos pelo

fabricante e específicos para essa plataforma de sequenciamento. Todas as corridas de sequenciamento foram feitas utilizando o MinION Mk1C® (Oxford Nanopore Technologies, Oxford, Reino Unido), e cada uma delas foi realizada por uma célula de fluxo (do inglês *flow cell*) MinION (versão FLO-MIN 106 R9; Mki 1 Spot-ON). Antes de cada corrida, a *flow cell* foi avaliada individualmente por sua quantidade de nanoporos ativos totais disponíveis para sequenciamento, de acordo com o protocolo do fabricante, sendo com isto utilizadas *flow cells* com valor igual ou maior que 800 poros disponíveis para sequenciamento.

4.2 ANÁLISE DE BIOINFORMÁTICA

O programa RAMPART da ARTIC Network foi utilizado simultaneamente à corrida de sequenciamento porque fornece em tempo real uma visão sobre a cobertura do genoma. Essa visualização é obtida através do alinhamento dos *reads* gerados ao genoma referência de SARS-CoV-2, obtido do GenBank (Genbank refseq NC_045512.2) (SAYERS et al., 2021) por meio do identificador MN908947. Essas informações são fornecidas por estatísticas e representações gráficas (AREVALO, et al., 2022). O protocolo de montagem dos genomas foi realizado pela empresa Mendelics Análise Genômica S.A. por meio de uma colaboração.

Posteriormente, os resultados foram obtidos da plataforma da empresa e foi realizada a análise de atribuição de variantes e linhagens, utilizando duas ferramentas. A primeira delas é o *Nextclade* (AKSAMENOV et al., 2021), e a outra é o *Pangolin* (O'TOOLE et al., 2021). Esses programas atribuem, através do genoma submetido, a classificação mais provável do vírus com base no conjunto de mutações, inserções e deleções no genoma (O'TOOLE et al., 2020).

As sequências fornecidas pela empresa foram selecionadas com base na cobertura do genoma e localização. Genomas com cobertura abaixo de 90%, e cobertura mínima de 50x foram selecionadas. Foi demarcado também o período de estudo, que compreende as datas de coleta de 4 de novembro de 2021 a 25 de abril de 2022, obtendo um total de 338 amostras (n = 338).

Mutações, deleções e inserções foram utilizadas para avaliar a dissimilaridade dos genomas usando o pacote R *vegan* v. 2.5.6 e o método do índice de Jaccard. A análise de escala multidimensional (MDS) e a distribuição das

variantes em Florianópolis foram plotadas usando o pacote R ggplot2 v. 4.2.0 (WICKHAM; NAVARRO; PEDERSEN, 2016).

4.3 ASPECTOS ÉTICOS

O presente estudo faz parte de um estudo intitulado "Estudo Colaborativo de Vigilância Genética de SARS-CoV-2 em cidades brasileiras" (CAAE: 51855621.4.2013.0121), que está aprovado no Comitê de Ética em Pesquisa com Seres Humanos (CEP/UFSC).

5 RESULTADOS

5.1 PERFIL DAS VARIANTES DE SARS-CoV-2 EM FLORIANÓPOLIS DURANTE O PERÍODO DE NOVEMBRO DE 2021 A ABRIL DE 2022

Foram analisados um total de 338 genomas completos de SARS-CoV-2, distribuídos em amostras coletadas entre o período de 4 de novembro de 2021 a 25 de abril de 2022 (Figura 4).

Nas amostras foram identificadas a VOC *Delta* (AY.101) e a VOC *Ômicron* (BA.1), além de 26 diferentes linhagens. O clado *Ômicron* foi identificado em 188 (55,6%) genomas, e o restante dos 150 genomas foi agrupado no clado *Delta* (44,1%).

Figura 4 - Amostragem total por período

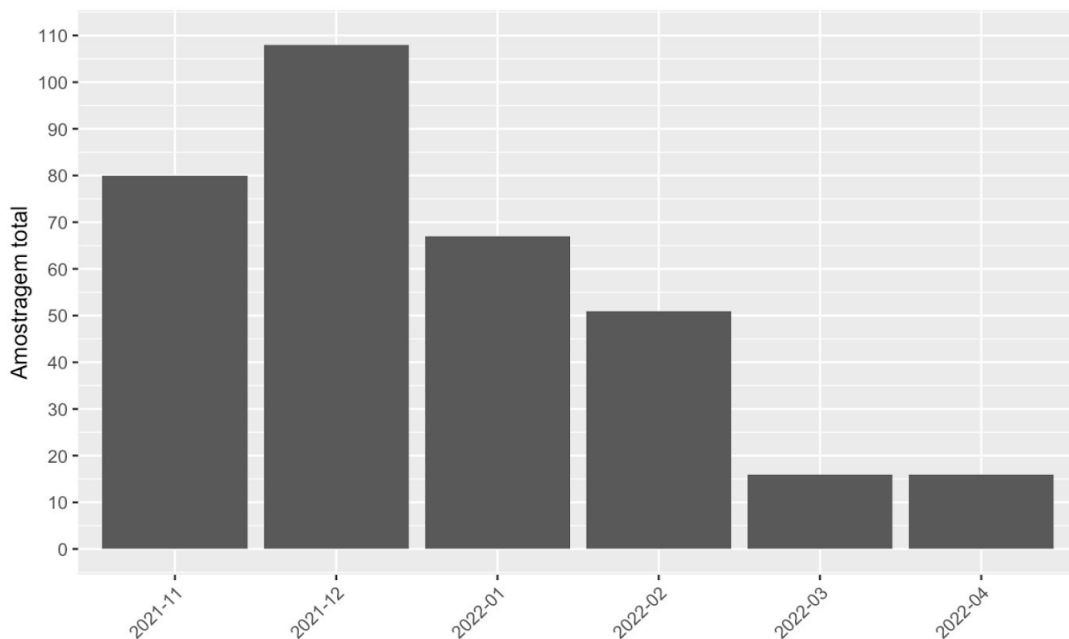


Figura 4 - Representação da quantidade de amostras obtidas mensalmente, de novembro de 2021 a abril de 2022.

5.2 MUDANÇA DO PERFIL DE VARIANTES (*DELTA* PARA *ÔMICRON*)

Mudanças no perfil das variantes foram analisadas mensalmente e de acordo com a frequência relativa (Figura 5). Foi analisada também a frequência relativa do perfil de linhagens das amostras obtidas (Figura 6). Este conjunto de dados foi analisado também proporcionalmente (Figura 7), de acordo com o número de amostras identificadas.

No total, foram identificadas 10 linhagens diferentes da VOC *Delta*, sendo que a AY.101 foi a mais frequente, tendo sido identificada em 63 das 150 amostras. Isto corresponde a 42% sobre o total de variante *Delta* e 18,6% sobre o total de amostras identificadas em todo o estudo.

Entre as diferentes linhagens da *Ômicron*, a BA.1.1 foi predominante no período estudado, representando 46% do total referente a esta variante e 20,7% sobre o total de amostras analisadas do estudo.

Figura 5 - Frequência Relativa por Variantes por mês

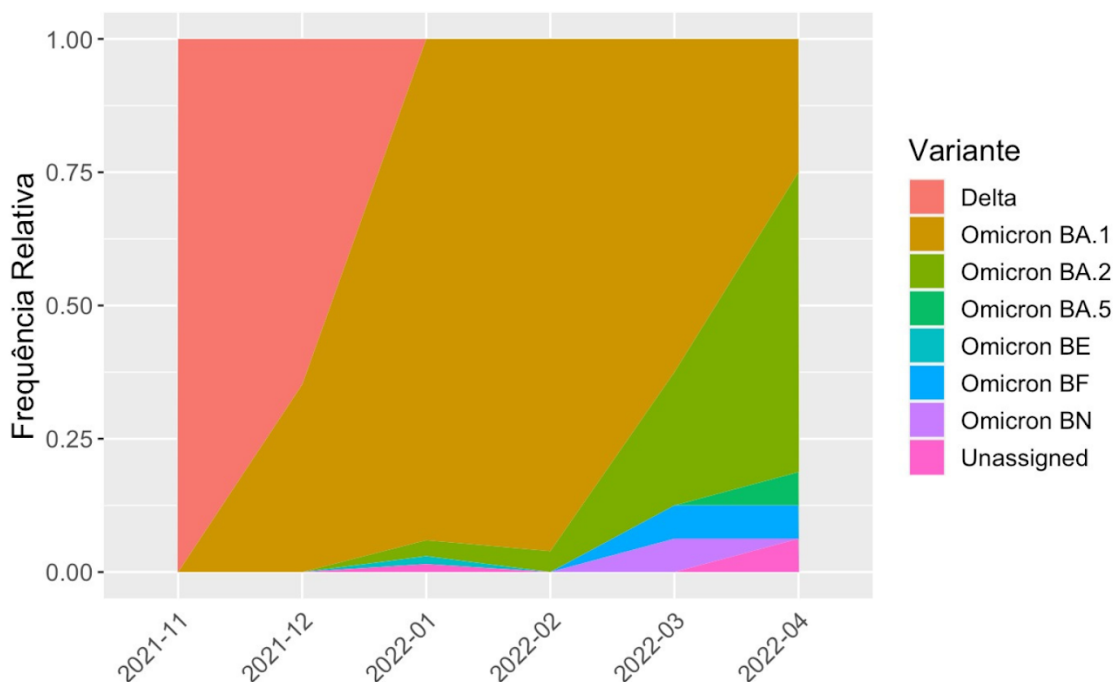


Figura 5 - Perfis de variantes de SARS-CoV-2 na região da Grande Florianópolis no período entre novembro de 2021 e abril de 2022.

Figura 6 - Frequência Relativa por Linhagens por mês

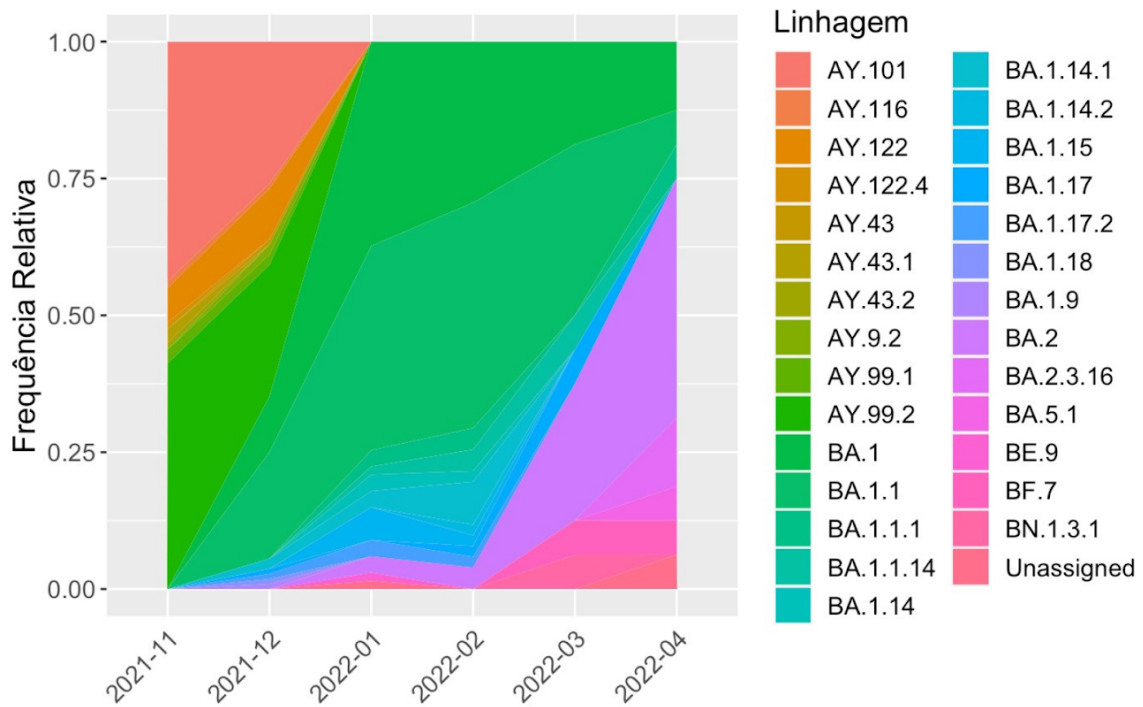


Figura 6 - Perfis de linhagens de SARS-CoV-2 na região da Grande Florianópolis no período entre novembro de 2021 e abril de 2022.

Os primeiros genomas sequenciados foram classificados como VOC *Delta*, representados principalmente pela linhagem AY.101 e AY.99.2, sendo que estes representaram a totalidade dos genomas identificados no mês de novembro. Nas amostras sequenciadas referentes ao mês de dezembro houve mais diversidade, sendo que de 108 genomas, 30 foram identificados como VOC *Ômicron*, incluindo 6 diferentes linhagens, entre elas a BA.1.

Em um mês após a introdução da VOC *Ômicron* BA.1, não houve mais nenhum genoma identificado pertencente ao clado *Delta*. A BA.1 foi predominante nos meses seguintes, representando mais de 90% das sequências identificadas em janeiro e fevereiro.

No mês de fevereiro houve a introdução da linhagem *Ômicron* BA.2, que aumentou gradualmente sua frequência até representar 56% das sequências identificadas no mês de abril. Entre os meses de março e abril houve a identificação de novas linhagens da *Ômicron*, BA.5, BE, BF e BN.

Figura 7 - Contagem da frequência de Variantes por mês

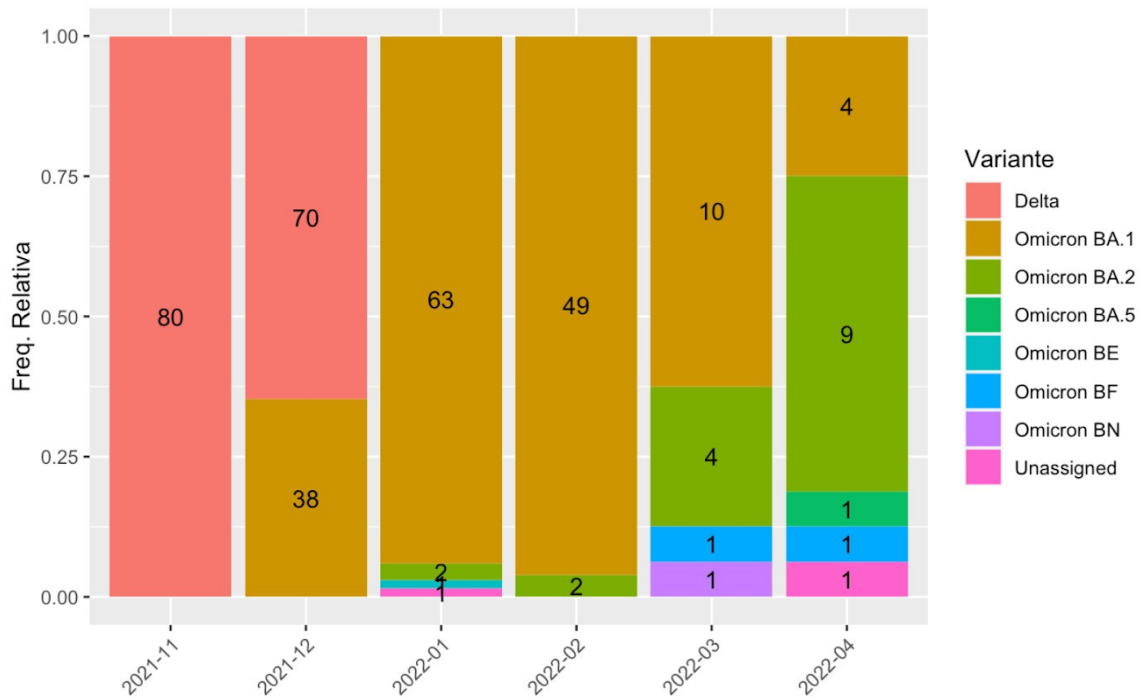


Figura 7 - Perfis de variantes de SARS-CoV-2 no período entre novembro de 2021 e abril de 2022 e de acordo com o número de amostras identificadas.

O gráfico de escala multidimensional (MDS) mostra o agrupamento dos genomas encontrados de SARS-CoV-2 de acordo com as mutações observadas. Foi possível observar 3 agrupamentos (Figura 8), entre as amostras identificadas como VOC *Delta* (AY.101); entre as identificações relacionadas à VOC *Ômicron* BA.1 e outro com as linhagens relacionadas à BA.2, BA.5, BE e BF. Também é possível ver a mesma relação entre linhagens (Figura 9).

Figura 8 - Gráfico de escala multidimensional (MDS) da matriz de dissimilaridade de Jaccard

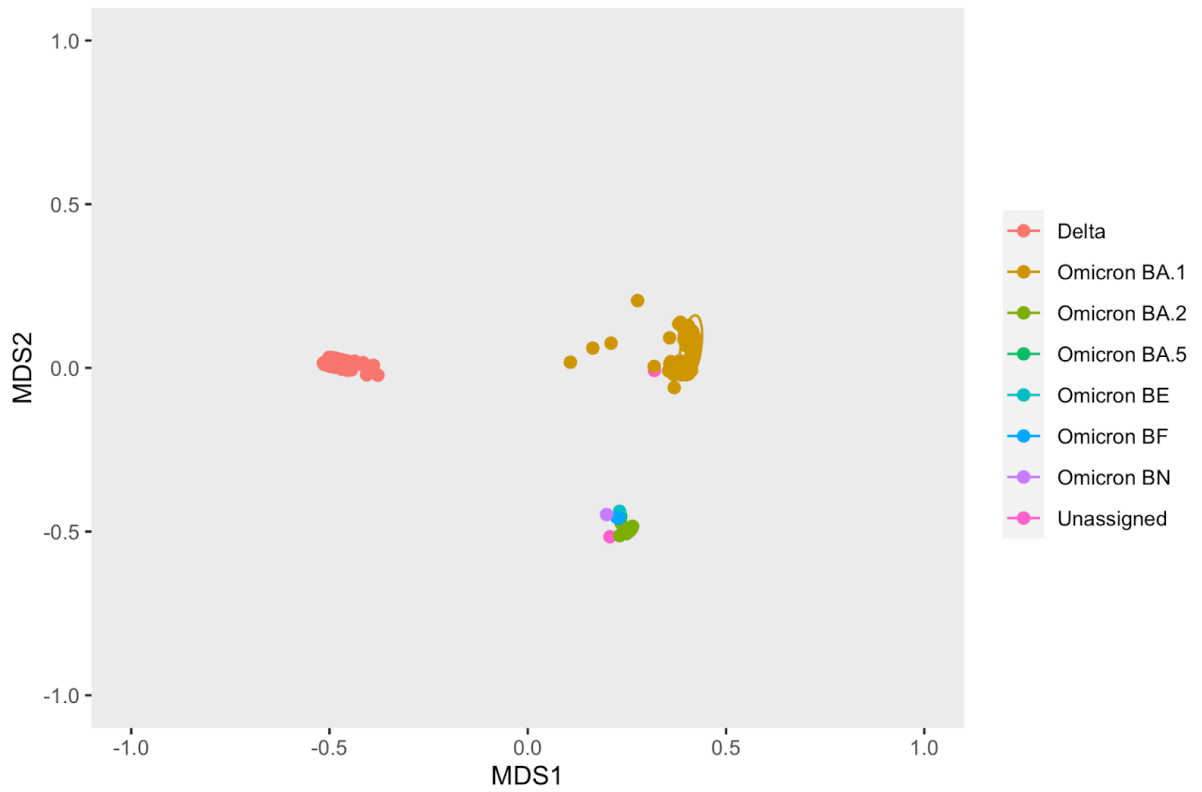


Figura 8 - Gráfico em escala multidimensional (MDS) da matriz de dissimilaridade de Jaccard entre genomas com base na presença e ausência de mutações, deleções e inserções, de acordo com variantes identificadas.

Figura 9 - Gráfico de escala multidimensional (MDS) da matriz de dissimilaridade de Jaccard

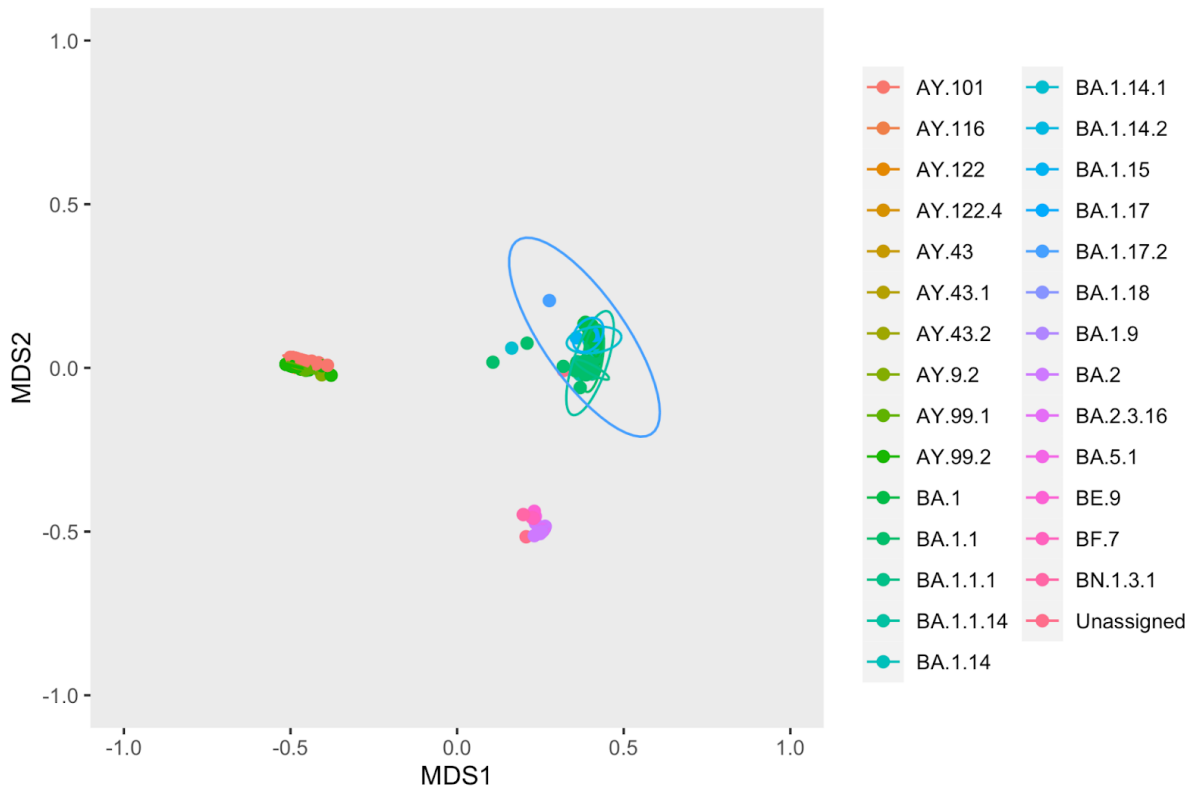


Figura 9 - Gráfico em escala multidimensional (MDS) da matriz de dissimilaridade de Jaccard entre genomas com base na presença e ausência de mutações, deleções e inserções, de acordo com linhagens identificadas.

6 DISCUSSÃO

Desde o início da pandemia causada por SARS-CoV-2, em 2020, ainda persistem desafios de saúde pública relacionados ao tratamento, diagnóstico, monitoramento em tempo real e vigilância genômica viral. Esta última configura-se como uma alternativa de prevenção de epidemias virais, e visa um conhecimento mais aprofundado sobre o patógeno por meio de estudos sobre sua evolução genética (WHO, 2021). Também é possível associar dados epidemiológicos aos estudos, permitindo avaliar a dinâmica de transmissão viral. Esta avaliação é mais eficaz quando unida a uma abordagem global.

Até junho de 2023, foram submetidas ao banco de dados GISAID mais de 230 mil sequências do genoma de SARS-CoV-2 provenientes do Brasil, representando 0,6% da cobertura sequencial dos casos reportados. Referente à VOC *Ômicron*, foram submetidas mais de 108 mil sequências de todo o país, sendo que destas mais de 6 mil foram oriundas do estado de Santa Catarina (GISAID, 2023).

Durante o curso da pandemia, esforços colaborativos utilizaram desta ferramenta para monitorar a evolução genética e identificar as variantes que impactaram globalmente na disseminação de SARS-CoV-2. A variante *Delta* foi identificada pela primeira vez na Índia (MLCOCHOVA et al., 2021), e as mutações nela observadas foram relacionadas a um aumento significativo na taxa de transmissão, superando todas as variantes previamente identificadas (BHATTACHARYA, et al., 2023).

Em Manaus, capital do estado do Amazonas, foi relatado o surgimento da variante *Gamma* em novembro de 2020 (FARIA et al., 2021). A transmissão desta variante foi considerável, além de ter sido observado que ocorreu sobretudo em centros urbanos mais densos das regiões Sudeste e Sul (GRAF et al., 2022). No estado de Santa Catarina, foi observado um aumento das taxas de mortalidade na ocasião da predominância das sub-linhagens P.1-*like* 2 desta variante (PADILHA et al., 2022).

Todas essas avaliações foram construídas a partir de dados de vigilância genômica. No estudo desenvolvido por Mahanta e colaboradores (2022) foi comparada a eficiência da estratégia em cada país do mundo. De 2020 até 2021 foi observado que a taxa de sequenciamento de SARS-CoV-2 aumentou, além da

submissão dos dados ao GISAID ter sido mais rápida. Com isso, os países em desenvolvimento melhoraram significativamente sua vigilância genômica. Dentre estes, o Brasil passou a sequenciar 0,9 para 5,6 genomas a cada mil casos, o que corresponde a um aumento de 6,5 vezes (MAHANTA; SABERWAL; SHARMA, 2022).

É possível perceber que há esforços reunidos para a melhora da eficiência de vigilância genômica. Porém, ainda há a necessidade de aprimoramento de tecnologias que se configuram como alternativas para abordá-la amplamente. Dentre estas, a plataforma MinION utiliza uma abordagem de sequenciamento por nanoporos da tecnologia ONT. Esta foi desenvolvida pela empresa Oxford Nanopore Technologies (Oxford, Reino Unido).

A plataforma de sequenciamento MinION se constitui como uma opção para o monitoramento de epidemias por suas características de portabilidade, redução do tempo de preparo de bibliotecas de sequenciamento, além de ser mais barata (AREVALO, et al., 2022; FREED et al. 2020). Entretanto, esta tecnologia ainda está em processo de consolidação devido aos fatores de qualidade envolvidos no sequenciamento.

Um dos parâmetros que foram associados à maior cobertura genômica foram os valores de Cq das amostras analisadas. Esse valor refere-se ao número de ciclos necessários para que a curva de amplificação ultrapasse a linha de *threshold*. Esse valor tem uma proporção inversa à quantidade de material genético na amostra e, conseqüentemente, à carga viral do paciente (PADILHA et al., 2022).

Os achados de Tshiabuila e colaboradores (2022) demonstram que na plataforma Illumina, tecnologia NGS, não é afetada por essa relação entre valores de cobertura genômica e Cq. No entanto, a tecnologia ONT teve resultado afetado por conta desta diferença. Procuramos eliminar este viés de nosso estudo, utilizando somente amostras com o valor de Cq <25, como mencionado anteriormente.

O presente estudo visou identificar sequências de amostras obtidas em Florianópolis no período de outubro de 2021 a abril de 2022, obtidas através do sequenciamento pela plataforma MinION (Oxford, Reino Unido). Este intervalo de tempo compreende a terceira onda da pandemia da COVID-19, marcado pelo aumento de casos pela VOC *Ômicron*.

Foi demonstrado anteriormente na pandemia que fatores como densidade populacional, presença de aeroportos internacionais e viagens sazonais estavam relacionados com a dispersão de linhagens pelo mundo e no Brasil (NICOLELIS, et al., 2021; LAMARCA, 2023). Foi observado também que regiões com essas características apresentaram taxas elevadas de transmissão viral no Brasil (GRAF et al., 2022). Isso ajuda a compreender a dinâmica de dispersão de SARS-CoV-2 que se manteve até o momento.

A cidade de Florianópolis, capital do estado de Santa Catarina, é característica por sua importância socioeconômica e turística, sendo a segunda cidade mais populosa do estado, com mais de 500 mil habitantes (IBGE, 2023). A presença de um aeroporto internacional na cidade também aumenta o fluxo de pessoas e de viagens sazonais, que são um fator de introdução de linhagens novas e de aumento da taxa de transmissão.

A variante *Ômicron* foi relatada mundialmente pela primeira vez em novembro de 2021, a partir de uma amostra obtida na África do Sul durante um surto de casos na região de Joanesburgo (VIANA, et al., 2022). No mesmo mês, a OMS a classificou como VOC (WHO, 2023) baseando-se nas mutações que ela apresentava em todo seu genoma e principalmente na proteína *Spike*. Estas foram relacionadas a maior transmissibilidade e grau de escape imunológico. A *Ômicron* acumula mais de 40 mutações, incluindo algumas que podem torná-la mais infecciosa, como D614G, N501Y e K417N (WEI et al., 2021). Outras podem aumentar a transmissão do vírus, como H655Y, N679K e P681H, e foram encontradas também nas VOC *Alfa* e *Delta* (ZHANG, et al., 2021).

Em comparação com a VOC *Delta*, a VOC *Ômicron* apresenta taxa de transmissibilidade 3 a 5 vezes maior, o que resultou em sua rápida disseminação em mais de 50 países em apenas um mês (GOWRISANKAR, et al., 2022). Essas alterações estão associadas à reinfecção e à redução da eficácia das vacinas. Além disso, a identificação da *Ômicron* coincidiu com um aumento na frequência de falhas na amplificação do gene S nos testes diagnósticos de PCR. Este fenômeno já tinha sido observado previamente com a variante *Alfa*, e foi atribuído a uma deleção de aminoácidos na proteína *Spike* (VIANA, et al., 2022).

É importante ressaltar que essas modificações não necessariamente implicaram em maior virulência, e nesse sentido, se caracterizam por se concentrarem nas vias aéreas superiores, em contraste com as variantes

previamente identificadas (CUI, et al., 2022). Este padrão ocorre devido à ligação da *Spike* à ACE2 ter se tornado mais estável, enquanto a fusão celular durante a infecção perdeu estabilidade (SUZUKI, et al., 2022).

Em outubro de 2021, os casos de detecção de SARS-CoV-2 no Brasil estavam em declínio, atingindo o menor patamar desde o início da pandemia. No mês seguinte, foi notificada a introdução da VOC *Ômicron* no estado de São Paulo. Em razão desta ocorrência, houve uma mudança no cenário pandêmico do Brasil, tanto que em janeiro de 2022 o Brasil atingiu o maior pico de casos notificados de toda a pandemia de COVID-19 (BRASIL, 2022). Notou-se que o número de mortos não seguiu esse aumento acentuado, sendo 75% menor do que o pico de mortes causado pela variante *Gamma* em 2021 (LAMARCA, 2023).

Os perfis genéticos encontrados neste estudo (Figura 5) permitiram a identificação da VOC *Delta* e *Ômicron*, além de 29 linhagens diferentes (Figura 6). Através da análise de frequência relativa, observamos que nos dois primeiros meses da terceira onda, a prevalência foi da VOC *Delta* AY.101, sendo rapidamente substituída pelo surgimento da VOC *Ômicron* BA.1. Posteriormente em fevereiro de 2022, ela foi substituída pela *Ômicron* BA.2.

Em Santa Catarina, estudos mostram a mesma relação entre a frequência observada. Foi possível notar uma dominância plena entre as linhagens da variante *Delta* até a primeira identificação de *Ômicron*, em dezembro de 2021. Na pesquisa de Padilha e colaboradores (2023) este padrão foi mostrado em uma análise feita entre dois grupos: a população geral e os trabalhadores da área da saúde. Na avaliação de ambos os perfis, foi demonstrada a importância deste segundo grupo como sentinela da tendência da frequência das linhagens observadas (PADILHA et al., 2023). Nossos dados estão de acordo com esse padrão observado, conforme evidenciado pela Figura 7, onde mostra que as 80 amostras são agrupadas na VOC *Delta*.

No período de dezembro, no estado de Santa Catarina, foram confirmados os primeiros casos da VOC *Ômicron* através do sequenciamento (SANTA CATARINA, 2022a). Na época, ela já representava dominância nos genomas identificados. Importante notar que, naquele mês, o estado replicou o padrão nacional de aumento do número de casos por dia, ocorrido após a emergência da variante *Ômicron* BA.1 (PADILHA et al., 2023).

Nas amostras obtidas no presente estudo, a representatividade de amostras identificadas como *Ômicron* foi de 35% no mês de Dezembro (Figura 8), e todas foram classificadas como relativas à linhagem BA.1. Este dado se relaciona com o encontrado por análise anterior, que mostra também que a cidade que apresentou maior prevalência de casos *Ômicron* foi Florianópolis e apontada como a provável introdução da variante no estado (SANTA CATARINA, 2022a). Vale ressaltar que foi observado que este perfil de mudança ocorreu mais rapidamente entre os profissionais de saúde do que na população em geral (PADILHA et al., 2023).

Ainda, no início de janeiro de 2022, o predomínio de amostras identificadas como *Ômicron* foi de mais de 95%, sendo que todas elas se concentravam na área da Grande Florianópolis (SANTA CATARINA, 2022b). Até esse momento, só haviam sido identificadas amostras da linhagem BA.1, porém no final do ano de 2021 identificou-se os primeiros casos da sublinhagem BA.1.1 (PADILHA et al., 2023).

Nosso estudo obteve resultados similares, em que 94% das amostras referentes à VOC *Ômicron* foram designadas como sendo da linhagem BA.1. Do restante, 2 amostras foram identificadas como sendo da linhagem BA.2, e 1 amostra como BE.

No final de janeiro de 2022, nas amostras sequenciadas no estado de Santa Catarina, havia domínio absoluto das linhagens relacionadas à variante *Ômicron*. Neste período houve uma diminuição no volume de amostras analisadas, que poderia estar relacionado à forma empregada para diagnóstico. Essa migração ocorreu das técnicas de RT-PCR para testes imunológicos rápidos, e isto pode ter impedido a captação de amostras para o monitoramento genômico pelas instituições de saúde (SANTA CATARINA, 2022c).

Ainda assim, dados da Fundação Oswaldo Cruz (FIOCRUZ) mostraram que naquele período a *Ômicron* foi detectada em mais de 95% dos genomas sequenciados em todo o território brasileiro. Isso é um demonstrativo de sua rápida dominância epidemiológica (FIOCRUZ, 2022).

No final de janeiro de 2022, Florianópolis era a cidade do Estado com a maior prevalência de casos pela VOC *Ômicron*, seguida por Joinville e São José. Na capital do Estado, as linhagens mais frequentes eram BA.1 e BA.1.1, com menor número de casos da linhagem BA.2 e outras BA (SANTA CATARINA, 2022c).

Em fevereiro, os dados de sequenciamento em Santa Catarina, obtidos tanto por Padilha e colaboradores (2023) quanto através deste estudo demonstraram uma

diminuição da prevalência de VOC *Ômicron* BA.1. Esta diminuição pode ser atribuída ao aumento da sublinhagem BA.1.1, que seguiu até o final de março, mês em que foi identificado aumento de casos da BA.2 (PADILHA et al., 2023).

No estudo realizado por Lamarca e colaboradores (2023), foram analisados dados de sequências dos genomas de SARS-CoV-2 depositados no GISAID durante a terceira onda da COVID-19. Nele, Santa Catarina é o segundo estado em que mais ocorreram importações da linhagem BA.2, com 45 ocorrências. Ele aponta também a dispersão interestadual, em que Santa Catarina recebeu 11 importações da linhagem BA.2 do estado de São Paulo (LAMARCA, 2023). Estes achados se relacionam com os fatores epidemiológicos citados anteriormente, como a densidade populacional, em que o estado de Santa Catarina está inserido.

Os resultados encontrados neste estudo são consistentes com um trabalho anterior que evidenciou que abril de 2022 havia predomínio quase pleno da VOC *Ômicron* BA.2 sobre as demais variantes nos genomas identificados (PADILHA et al., 2023). Neste aspecto, nosso estudo identificou adicionalmente as sublinhagens BE, BF e BN do clado *Ômicron*.

Além disso, dos 338 genomas analisados, foi observado que em somente uma das amostras não houve identificação de linhagem, somente da variante, nomeada como *Unassigned* (Figura 7). Isto significa que o *software* utilizado não conseguiu atribuir a classificação de linhagem de maneira inequívoca. Então, ao invés de fornecer um resultado falso positivo ele retorna o resultado como *Unassigned*, ou seja: não classificado.

Isto pode ser atribuído a diversos fatores, entre eles a técnica de amplificação utilizada. Ela foi estabelecida antes da primeira detecção da VOC *Ômicron* em 2021, e baseia-se na amplificação da totalidade do genoma viral. No caso das amostras *Ômicron*, a qualidade do sequenciamento sofre influência das 50 mutações que ela carrega no genoma. Devido a esta mudança de cenário, a técnica teve que ser otimizada posteriormente.

7 CONCLUSÃO

A partir da realização deste projeto, foi possível identificar a mudança do perfil de variantes no contexto da terceira onda da COVID-19 em Florianópolis. Observou-se que a VOC *Delta* foi substituída pela VOC *Ômicron* BA.1 em dezembro de 2021, posteriormente substituída pela VOC *Ômicron* BA.2 entre março e abril de 2022. Hoje, a tecnologia ONT é conhecida por sua rapidez na geração de dados, além de sua possível portabilidade e menor custo. Mesmo com os desafios envolvendo esta técnica, neste estudo foram encontradas similaridades no perfil das variantes comparando com tecnologias NGS (hoje o padrão ouro de sequenciamento para diagnóstico), considerando o mesmo período e região. Além disso, é importante ressaltar a relevância da vigilância genômica como ferramenta de monitoramento e prevenção doenças infecto-contagiosas que possam levar ao surgimento de epidemias e pandemias como a vivida nos últimos tempos.

REFERÊNCIAS

AKSAMENOV, Ivan et al. Nextclade: clade assignment, mutation calling and quality control for viral genomes. **Journal of Open Source Software**. Vol 6.67 Setembro de 2021.

ALEEM, Abdul, Akbar Samad AB, Vaqar S. Emerging Variants of SARS-CoV-2 and Novel Therapeutics Against Coronavirus (COVID-19). **StatPearls**. 8 de maio de 2023. PMID: 34033342

ARÉVALO, Maria T. et al. A Rapid, Whole Genome Sequencing Assay for Detection and Characterization of Novel Coronavirus (SARS-CoV-2) Clinical Specimens Using Nanopore Sequencing. **Frontiers Microbiology**. Vol 13:910955. Junho de 2022.

ARTIC NETWORK. Real-Time molecular epidemiology for outbreak response. 2023. Disponível em <<https://artic.network/>> Acesso em 21 de maio de 2023.

ARTIC NETWORK. SARS-CoV-2. 2020. Disponível em <<https://artic.network/ncov-2019>> Acesso em 21 de maio de 2023.

BARSKI, A. et al. Nature Milestones: Genomic Sequencing. Milestone 6. 2021 Disponível em: <https://www.nature.com/collections/genomic-sequencing-milestones>. Acesso em 31 de maio de 2021.

BHATTACHARYA, Manojit et al. Delta variant (B.1.617.2) of SARS-CoV-2: current understanding of infection, transmission, immune escape, and mutational landscape. **Folia Microbiologica**, [s. l], v. 68, n. 1, p. 17-28, 12 ago. 2022.

BRANT, Ayslan Castro et al. SARS-CoV-2: from its discovery to genome structure, transcription, and replication. **Cell Bioscience**. Vol 11, 136. Julho de 2021

BRASIL. Ministério da Saúde. Painel de casos de doença pelo coronavírus 2019 (COVID-19) no Brasil pelo Ministério da Saúde. Brasília, v.2.0. Novembro/2022. Disponível em <<https://covid.saude.gov.br/>> Acesso em 21 de maio de 2023.

BULL, Rowena A et al. Analytical validity of nanopore sequencing for rapid SARS-CoV-2 genome analysis. **Nature Communications** Vol. 11, 6272. Dezembro de 2020.

CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC). COVID-19, Variants and Genomic Surveillance for SARS-CoV-2. **2022**. Disponível em <<https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-surveillance.html>> Acesso em 10 de maio de 2023.

CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC). SARS-CoV-2 Variant Classifications and Definitions. **2022**. Disponível em <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html#anchor_1632158885160> Acesso em 10 de maio de 2023.

CHAND, Meera et al. Investigation of novel SARS-COV-2 variant Variant of Concern 202012 / 01 Detection of an epidemiological cluster associated with a new variant of concern Nomenclature of variants in the UK Current epidemiological findings. **Public Health England**, [S. I.], n. p. 1–11, Dezembro 2020

CHEN, Bin. et al. Overview of lethal human coronaviruses. **Signal transduction and targeted therapy**. Vol 5.1 Junho de 2020.

CORONAVIRIDAE STUDY GROUP OF THE INTERNATIONAL COMMITTEE ON TAXONOMY OF VIRUSES. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. **Nature Microbiology** Vol. 5, pag. 536–544. Abril de 2022.

CUI, Jie; LI, Fang; SHI, Zheng-Li. Origin and evolution of pathogenic coronaviruses. **Nature Reviews Microbiology**, [S.L.], Vol. 17, n. 3, p. 181-192, 10 dez. 2018. Dashboard. 2022. Disponível em: <<https://covid19.who.int>>

DANECEK, Petr, et al. Twelve years of SAMtools and BCFtools, **GigaScience**, Vol. 10, n. 2. Fevereiro de 2021.

DAVIES, Nicholas G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. **Science**. Vol 372. 2021

DEAMER, David et al. Three decades of nanopore sequencing. **Nature biotechnology**. Vol 34(5), pag 518-524. Maio de 2016.

FARIA, Nuno R. et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. **Science**, [S. I.], v. 372, n. 6544, p. 815–821, Abril 2021.

FOOD AND DRUG ADMINISTRATION (FDA). SARS-CoV-2 Viral Mutations: Impact on COVID-19 Tests. 2022. Disponível em: <<https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/sars-cov-2-viral-mutations-impact-covid-19-tests>> Acesso em 30 de maio de 2023.

FREED, Nikki E et al. Rapid and inexpensive whole-genome sequencing of SARS-CoV-2 using 1200 bp tiled amplicons and Oxford Nanopore Rapid Barcoding. **Biology Methods and Protocols** Vol 5. Julho de 2020.

GHEDIN, Elodie; et al. Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. **Nature**, Vol. 437, p. 1162-1166, out. 2005.

GIANI, Alice Maria et al. Long walk to genomics: history and current approaches to genome sequencing and assembly. **Computational And Structural Biotechnology Journal**, [S.L.], v. 18, p. 9-19, 2020.

GOWRISANKAR, A. et al. Omicron: a mysterious variant of concern. The European **Physical Journal Plus**, [s. l], v. 137, n. 1, p. 137-145, jan. 2022.

GRAF, Tiago et al. Phylogenetic-based inference reveals distinct transmission dynamics of SARS-CoV-2 lineages Gamma and P.2 in Brazil. **Iscience**. Vol.25. Abril de 2022.

GRUBAUGH, Nathan D. et al. Making Sense of Mutation: what D614G means for the covid-19 pandemic remains unclear. **Cell**, [s. l], v. 182, n. 4, p. 794-795, ago. 2020.

HAAGMANS, Bart L et al. Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. **The Lancet Infectious Diseases**, [s. l], v. 14, n. 2, p. 140-145, fev. 2014.

IBGE - Instituto Brasileiro de Geografia e Estatística. Panorama de Florianópolis, Santa Catarina, Brasil. V4.6.42. Disponível em: <https://cidades.ibge.gov.br/brasil/sc/florianopolis/panorama>. Acesso em 20 de maio de 2023.

INTERNATIONAL COMMITTEE ON TAXONOMY OF VIRUSES (ICTV). Coronaviridae, Family: Coronaviridae. 2022. Disponível em: <https://ictv.global/report_9th/RNAspos/Nidovirales/Coronaviridae> Acesso em 05 de maio de 2023.

ISLAM, Md. Jahirul et al. A review on structural, non-structural, and accessory proteins of SARS-CoV-2: highlighting drug target sites. **Immunobiology**, [s. l], Vol. 228, n. 1, p. 152302, jan. 2023.

JESUS, Jaqueline Goes de et al. Importation and early local transmission of COVID-19 in Brazil, 2020. **Revista do Instituto de Medicina Tropical de São Paulo**, [s. l], v. 62, p. 137-167, 2020.

JESUS, Jaqueline Goes de, et al. Acute Vector-Borne Viral Infection: Zika and MinION Surveillance. **ASM Journals**. Vol. 7, No. 4. Agosto de 2019.

KANNAN, Shantani; SUBBARAM, Kannan; ALI, Sheeza; KANNAN, Hemalatha. Molecular Characterization and Amino Acid Homology of Nucleocapsid (N) Protein in SARS-CoV-1, SARS-CoV-2, MERS-CoV, and Bat Coronavirus. **Journal Of Pure And Applied Microbiology**, [S.L.], v. 14, n. 1, p. 757-763, 23 maio 2020.

KARIM, Salim S. Abdool et al. New SARS-CoV-2 Variants — Clinical, Public Health, and Vaccine Implications. **New England Journal Of Medicine**, [S.L.], v. 384, n. 19, p. 1866-1868, 13 maio 2021.

KHAILANY, Rozhgar et al. Genomic characterization of a novel SARS-CoV-2. **Gene Reports**, Volume 19. Junho de 2020.

KORBER, Bette et al. Tracking Changes in SARS-CoV-2 Spike: evidence that d614g increases infectivity of the covid-19 virus. **Cell**, [s. l], v. 182, n. 4, p. 812-827, ago. 2020.

LAMARCA, Alessandra P. et al. The Omicron Lineages BA.1 and BA.2 (Betacoronavirus SARS-CoV-2) Have Repeatedly Entered Brazil through a Single Dispersal Hub. **Viruses**, [S.L.], v. 15, n. 4, p. 888, 30 mar. 2023.

MAHANTA, Utkarsha et al. Are Countries Becoming Better at SARS-CoV-2 Genomic Surveillance?. **Frontiers in Public Health**. Vol. 10. Abril de 2022.

NICOLELIS, Miguel A. L. et al. The impact of super-spreader cities, highways, and intensive care availability in the early stages of the COVID-19 epidemic in Brazil. **Scientific Reports**, [s. l], v. 11, n. 1, p. 137-145, 21 jun. 2021.

O'TOOLE, Áine et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. **Virus Evolution**. Vol 7(2):. Julho de 2021.

PADILHA, Dayane et al., Emergence of Two Distinct SARS-CoV-2 Gamma Variants and the Rapid Spread of P.1-like-II SARS-CoV-2 during the Second Wave of COVID-19 in Santa Catarina, Southern Brazil. **Viruses**. Vol. 14, p.695. Março de 2022.

PADILHA, Dayane Azevedo *et al.* Genomic Surveillance of SARS-CoV-2 in Healthcare Workers: a critical sentinel group for monitoring the SARS-CoV-2 variant shift. **Viruses**, [S.L.], v. 15, n. 4, p. 984-992, 17 abr. 2023.

PECK, Kayla M.; LAURING, Adam S. Complexities of Viral Mutation Rates. **Journal of Virology**. Vol. 92, n. 14, p. 1–8, 2018.

PETRACKOVA, Anna et al. Standardization of Sequencing Coverage Depth in NGS: recommendation for detection of clonal and subclonal mutations in cancer diagnostics. **Frontiers In Oncology**, [s. l], v. 9, p. 137-827, 4 set. 2019.

QUICK, Joshua et al. Real-time, portable genome sequencing for Ebola surveillance. **Nature**. Volume 530, páginas 228–232. Fevereiro de 2016.

RAMBAUT Andrew et al. A Dynamic Nomenclature Proposal for SARS-CoV-2 Lineages to Assist Genomic Epidemiology. **Nature Microbiology**, Vol. 5. Julho de 2020.

SANCHES, Paulo R.s. et al. Recent advances in SARS-CoV-2 Spike protein and RBD mutations comparison between new variants Alpha (B.1.1.7, United Kingdom), Beta (B.1.351, South Africa), Gamma (P.1, Brazil) and Delta (B.1.617.2, India). **Journal Of Virus Eradication**, [s. l], v. 7, n. 3, p. 100054, set. 2021.

SANTA CATARINA, Secretaria de Estado da Saúde de Santa Catarina. Portaria SES nº 1173, de 22 de outubro de 2021. Institui a Rede de Vigilância Genômica do vírus SARS-CoV-2 no Estado de Santa Catarina. 2021a. Publicado no **Diário Oficial do Estado de Santa Catarina**. Disponível em: <<https://www.legisweb.com.br/legislacao/?id=422057>>. Acesso em 30 de maio de 2023.

SANTA CATARINA, Secretaria de Estado da Saúde de Santa Catarina. Vigilância Genômica do SARS-CoV-2: Dados até a Semana Epidemiológica 48/2021. 2021b. Disponível em: http://lacen.saude.sc.gov.br/arquivos/boletim_7-genomica-lacen-dive-48.pdf. Acesso em 24 de maio de 2023.

SANTA CATARINA, Secretaria de Estado da Saúde de Santa Catarina. Vigilância Genômica do SARS-CoV-2: Dados até a Semana Epidemiológica 50/2021. 2022a. Disponível em: http://lacen.saude.sc.gov.br/arquivos/boletim_8-genomica-lacen-dive-50.pdf. Acesso em 28 de maio de 2023.

SANTA CATARINA, Secretaria de Estado da Saúde de Santa Catarina. Vigilância Genômica do SARS-CoV-2: Dados até a Semana Epidemiológica Nº 04. 2022b. Disponível em: http://lacen.saude.sc.gov.br/arquivos/boletim_genomica-lacen-dive-semana04-2022.pdf. Acesso em 24 de maio de 2023.

SANTA CATARINA, Secretaria de Estado da Saúde de Santa Catarina. Vigilância Genômica do SARS-CoV-2: Dados até a Semana Epidemiológica Nº 09. 2022c Disponível em: http://lacen.saude.sc.gov.br/arquivos/Boletim_Genomico_semana_09_2022.pdf. Acesso em 24 de maio de 2023.

SANTA CATARINA, Secretaria de Estado da Saúde de Santa Catarina. Vigilância Genômica do SARS-CoV-2: Dados até a Semana Epidemiológica Nº 20. 2022d Disponível em: http://lacen.saude.sc.gov.br/arquivos/Boletim_Genomico_Semana_20_2022.pdf. Acesso em 30 de maio de 2023.

SANTA CATARINA, Secretaria de Estado da Saúde de Santa Catarina. Vigilância Genômica do SARS-CoV-2: Dados até a Semana Epidemiológica Nº 02/2023. 2023. Disponível em: http://lacen.saude.sc.gov.br/arquivos/Boletim_Genomico_Semana_20_2022.pdf. Acesso em 30 de maio de 2023.

SAYERS, Eric W et al. GenBank. **Nucleic Acids Research**. Volume 49, Pág D92–D96. Janeiro de 2021.

SCHLEMPER JUNIOR, Bruno R. S., Covid-19 a Guerra da Desinformação. O Desafio da Ciência, Bioética e Socialização do Saber. 1ª Edição. **Editores Appris**. Outubro de 2022.

SHEN, Wei et al. SeqKit: a cross-platform and ultrafast toolkit for fasta/q file manipulation. **Plos One**, [S.L.], v. 11, n. 10, p. 134-145, 5 out. 2016.

SHU, Y. L.; MCCAULEY, J. GISAID: global initiative on sharing all influenza data - from vision to reality. **Eurosurveillance**, v. 22, p. 2–4, Março de 2020.

SIMS, David et al. Sequencing depth and coverage: key considerations in genomic analyses. **Nature Reviews Genetics**. Vol. 15, pág 121–132. Janeiro de 2014.

SLAVOV, Svetoslav, 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**. Vol. 93, pág 6782-6787. Dez 2021.

TSHIABUULA, Derek *et al.* Comparison of SARS-CoV-2 sequencing using the ONT GridION and the Illumina MiSeq. **Bmc Genomics**, [S.L.], v. 23, n. 1, p. 327-345, 22 abr. 2022.

VALVERDE, R. Omicron variant accounts for more than 95% of the genomes sequenced in Brazil. **Agência de Notícias Fiocruz**, 2022. Disponível em <<https://portal.fiocruz.br/en/news/omicron-variant-accounts-more-95-genomes-sequenced-brazil>> Acesso em 30 de maio de 2023.

VIANA Raquel et al. Rapid Epidemic Expansion of the SARS-CoV-2 Omicron Variant in Southern Africa. **Nature**. Vol. 603. Janeiro de 2022.

VOLZ, Erik et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. **Cell**. Volume 184, pag 64-75. Janeiro de 2021.

WANG, Yunhao et al. Nanopore sequencing technology, bioinformatics and applications. **Nature Biotechnology**. Volume 39, 1348–1365. Novembro de 2021.

WEI, Changshuo et al. Evidence for a mouse origin of the SARS-CoV-2 Omicron variant. **Journal Of Genetics And Genomics**, [s. l], v. 48, n. 12, p. 1111-1121, dez. 2021.

WICKHAM, H. Ggplot2: Elegant Graphics for Data Analysis; **Springer**: New York, NY, EUA, 2016. Disponível em: <https://ggplot2.tidyverse.org>. Acesso em 30 de maio de 2023.

WORLD HEALTH ORGANIZATION (WHO). Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health. 2021. Disponível em: <https://www.who.int/publications/i/item/9789240018440> Acesso em 20 de maio de 2023.

WORLD HEALTH ORGANIZATION (WHO). Statement on the update of WHO's working definitions and tracking system for SARS-CoV-2 variants of concern and variants of interest. 2023. Disponível em: <https://www.who.int/news/item/16-03-2023-statement-on-the-update-of-who-s-working-definitions-and-tracking-system-for-sars-cov-2-variants-of-concern-and-variants-of-interest> Acesso em 2 de junho de 2023.

WORLD HEALTH ORGANIZATION (WHO). Novel Coronavirus(2019-nCoV), Situation Report – 22. 2020. Disponível em: <<https://www.who.int/docs/default-source/coronaviruse/situation-reports.pdf>> Acesso em 15 de maio de 2023.

WORLD HEALTH ORGANIZATION (WHO). Tracking SARS-CoV-2 variants. 2022. Disponível em: <<https://www.who.int/activities/tracking-SARS-CoV-2-variants>> Acesso em 21 de maio de 2023.

WU, Aiping et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. **Cell Host & Microbe**, [s. l], v. 27, n. 3, p. 325-328, mar. 2020.

YE, C., et al. Exploiting sparseness in de novo genome assembly. **Bmc Bioinformatics**, v. 13, n. 6, p.1-8, 19 abr. 2012. Springer Nature.

ZHU, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. **New England Journal of Medicine**. Vol. 382, p.727-733. Fevereiro de 2020.