

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA**

FLAVIA CRISTINA DA SILVA SALES

**Análise genômica de SARS-CoV-2 na cidade de São Caetano do Sul,
São Paulo, Brasil.**

São Paulo

2023

Flavia Cristina da Silva Sales

Análise genômica de SARS-CoV-2 na cidade de São Caetano do Sul, São Paulo, Brasil.

Versão corrigida. Resolução CoPGr 6018/11, de 01 de novembro de 2011. A versão original está disponível na Biblioteca FMUSP.

Tese apresentada à Faculdade de Medicina da Universidade de São Paulo para obtenção do título de Doutor em Ciências Programa de Doenças Infecciosas e Parasitárias
Área de concentração: Doenças Infecciosas e Parasitárias
Orientadora: Profa. Dra. Ester Cerdeira Sabino

São Paulo

2023

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Sales, Flávia Cristina da Silva
Análise genômica de SARS-CoV-2 na cidade de São
Caetano do Sul, São Paulo, Brasil / Flávia Cristina
da Silva Sales. -- São Paulo, 2023.
Tese(doutorado)--Faculdade de Medicina da
Universidade de São Paulo.
Programa de Doenças Infecciosas e Parasitárias.
Orientadora: Ester Cerdeira Sabino.

Descritores: 1.SARS-CoV-2 2.Epidemiologia
genômica 3.Atenção primária 4.Doenças infecciosas
5.COVID-19

USP/FM/DBD-131/23

Responsável: Erinalva da Conceição Batista, CRB-8 6755

SALES, Flavia Cristina da Silva

Título: Análise genômica de SARS-CoV-2 na cidade de São Caetano do Sul, São Paulo, Brasil.

Tese apresentada à Faculdade de Medicina da Universidade de São Paulo para obtenção do título de Doutor em Ciências do Programa de Doenças Infecciosas e Parasitárias; Área de concentração: Doenças Infecciosas e Parasitárias.

Aprovado em: 27 de Abril de 2023.

Banca examinadora

Prof(a) Dr.(a) Silvia Figueiredo Costa

Instituição: Faculdade de Medicina da Universidade de São Paulo - USP

Prof(a) Dr.(a) José Luiz Proença Modena

Instituição: Universidade Estadual de Campinas - UNICAMP

Prof(a) Dr.(a) Ana Carolina Gomes Jardim

Instituição: Universidade de Uberlândia - UFU

"Afinal, o que é um cientista, então? Ele é um homem curioso que olha através dos olhos de uma fechadura, a fechadura da natureza, tentando saber o que está acontecendo"

Jacques Yves Cousteau

DEDICATÓRIA

Início essa dedicatória, à minha orientadora, Dra. Ester Cerdeira Sabino, pela inspiração e por todas as oportunidades incríveis que me levaram até aqui. Através da sua confiança, tive o prazer de conhecer e trabalhar com pessoas que tratam a ciência com seriedade e paixão.

Ao Dr. Nuno Faria, da Universidade de Oxford/Imperial College, pelos diversos *insights* e sugestões a esse trabalho, que mente brilhante!

Aos pesquisadores do Grupo CADDE pela oportunidade, aprendizado, viagens, horas e horas trabalhadas e momentos únicos compartilhados. Que profissionais incríveis! Todos foram essenciais em diversas etapas desse trabalho e de toda a minha trajetória: Filipe Romero, Carlos Prete, Leandro Abade, Lewis Buss, Jaqueline Goes, Camila Maia, Lucas Franco, Geovana Pereira, Ian Nunes, Thais Coletti, Pamela Andrade, Esmenia Coelho, Midiã Ferreira, Franciane Mendes, Melina Mosquera, Valquiria Reis, Bianca Costa, Raissa Heloisa Eliodoro, Mayara Bertanhe e especialmente, aos meus queridos amigos Darlan Candido, Ingra Claro, Mariana Severo e Erika Regina Manulli (Tudo vai dar certo!) pelo apoio, amizade, risadas, abraços e choros compartilhados.

Aos colegas do Laboratório de Parasitologia LIM 46, Léa Campos Oliveira, Cecilia Salete, Fabio Ghilardi, Carlos Henrique Valente, Marcia Hage, Natália Souza de Godoy, Regina Maia, obrigada por compartilhar o conhecimento de vocês.

Um agradecimento especial, às minhas queridas amigas que terminaram seu ciclo no Instituto, porém, a amizade permaneceu fora dele, Roberta Ruedas, Natalia Bueno, Camila Cruz, Michelli Medeiros e Marcela Souza.

A todos do grupo Zibra 2 (Dra. Marta Giovanetti, Joilson Xavier), especialmente, ao Prof. Luiz Alcantara pelas oportunidades, experiências e diversas viagens pelo Brasil, treinando, compartilhando e fazendo ciência.

Ao grupo do Laboratório de Virologia LIM 52, Profa. Dra. Cassia Mendes, Tania Regina, Camila Romano, Anderson de Paula, Noely Evangelista e Heuder Paião e ao grupo do Laboratório LIM 54, Professora Silvia Costa, Ana Paula Marchi, Alessandra Luna, Camila Rizek pela amizade.

A todos os profissionais do grupo da Plataforma Corona São Caetano e USCS, que realizaram um trabalho excepcional durante uma das fases mais difíceis da história da humanidade.

Aos grupos de pesquisa que disponibilizaram as sequências produzidas de SARS-CoV-2 no GISAID.

Dedico este trabalho aos meus queridos pais (Marineth Lima e Francisco de Sales) e irmãos (Felipe Sales e Felix Sales) e ao meu amor (Guilherme Vulcano) por todo suporte, incentivo, compreensão e carinho.

Finalmente, me solidarizo e deixo minhas profundas condolências a todas as pessoas que tiveram seus entes queridos perdidos e que sofreram com as consequências desse período difícil da nossa história.

Gratidão!

Esta dissertação está de acordo com as seguintes normas em vigor no momento desta publicação:

Referências: adaptado de International Committee of Medical Journals Editor (Vancouver)

Universidade de São Paulo. Agência USP de Gestão da Informação Acadêmica. Diretrizes para apresentação de dissertações e teses da USP: parte I (ABNT) / Agência USP de Gestão de Informação Acadêmica; Vânia Martins Bueno de Oliveira Funaro, coordenadora; Vânia Martins Bueno de Oliveira Funaro. [et al]. – 4ª Ed. – São Paulo: AGUIA, 2020.

AGRADECIMENTOS

O presente trabalho foi realizado com o apoio da Faculdade de Medicina da Universidade de São Paulo, do Instituto de Medicina Tropical da Universidade de São Paulo e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) através dos seguintes processos: nº 2018/14389-0 e 2018/25468-9.

RESUMO

Sales, FCS. Vigilância genômica de SARS-CoV-2 na cidade de São Caetano do Sul, São Paulo, Brasil. [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2023.

O mundo testemunhou inúmeras epidemias e pandemias que afetaram milhões de vidas. Apesar de nossos avanços em medicina e pesquisa, continuamos a ser desafiados com novos patógenos que representam uma ameaça à vida humana, à segurança econômica global e ao sistema de saúde. Avanços recentes permitiram que os genomas do Coronavírus da Síndrome Respiratória Aguda Grave 2 (SARS-CoV-2) – o agente causador do COVID-19 – pudessem ser identificados e caracterizados rapidamente. O estudo de sequências genômicas aliadas a um conjunto de metadados robustos e análises epidemiológicas tradicionais podem contribuir para o entendimento da transmissão e controle de uma doença infecciosa. Neste estudo, utilizamos dados epidemiológicos e genômicos coletados de um detalhado banco de dados originados da Plataforma Corona São Caetano, um programa de cuidados primários desenvolvido para atender pacientes com suspeita de COVID-19, centralizando dados clínicos e amostrais. Usamos essa infraestrutura para descrever em detalhes a evolução das variantes do SARS-CoV-2 no município de São Caetano do Sul, uma cidade intensamente conurbada com a cidade de São Paulo, durante o primeiro ano da epidemia. Um total de 1063 amostras positivas para SARS-CoV-2 foram selecionadas e dessas 879 sequências genômicas com cobertura acima de 75x foram geradas. Demonstramos a persistência das linhagens B.1.1.28 e B.1.3.3, início da circulação das VOIs Zeta (P.2) e múltiplas introduções da VOC Gama (P.1) no município. Destacamos também, a incidência dos casos na comunidade, identificando as regiões do município mais afetadas e correlacionamos as informações clínicas com as variantes locais. Nossos dados sugerem que sintomas relacionados à doença mais grave, como dispnéia e febre prolongada, foram aproximadamente duas vezes mais comuns entre Gama/VOC e Zeta/VOI em comparação com as cepas ancestrais. Ageusia e anosmia foram menos comuns na linhagem Gama e as manifestações neurológicas foram mais comuns no Zeta/VOI. A taxa de letalidade e a taxa de hospitalização foram maiores durante o período da circulação da Gama/VOC. Nossos dados sugerem que Gama/VOC e Zeta/VOI

causaram doença mais grave e isso pode ter contribuído para o maior nível de hospitalização e óbitos durante a segunda onda no Brasil. Além disso, destacamos como estudos de epidemiologia genômica aliados a uma amostragem representativa e dados epidemiológicos detalhados são fundamentais para oferecer *insights* sólidos sobre a transmissão do COVID-19.

Descritores: SARS-CoV-2, epidemiologia genômica, atenção primária, doenças infecciosas, COVID-19.

ABSTRACT

Sales, FCS. Genomic analysis of SARS-CoV-2 in the city of São Caetano do Sul, São Paulo, Brazil. [thesis]. São Paulo: Faculty of Medicine, University of São Paulo; 2023.

The world has witnessed countless epidemics and pandemics that have affected millions of lives. Despite our advances in medicine and research, we continue to be challenged with new pathogens that pose a threat to human life, global economic security and the health system. Recent advances have allowed the genomes of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) – the causative agent of COVID-19 – to be rapidly identified and characterized. The study of genomic sequences combined with a set of robust metadata and traditional epidemiological analyzes can contribute to the understanding of the transmission and control of an infectious disease. In this study, we used epidemiological and genomic data collected from a detailed database originating from the Plataforma Corona São Caetano, a primary care program developed to care for patients with suspected COVID-19, centralizing clinical and sample data. We used this infrastructure to describe in detail the evolution of SARS-CoV-2 variants in the municipality of São Caetano do Sul, a city strongly conurbated with the city of São Paulo, during the first year of the epidemic. A total of 1063 positive selections for SARS-CoV-2 were selected and from these 879 good quality genomic sequences were generated. We demonstrate the persistence of the B.1.1.28 and B.1.3.3 lineages, the beginning of the circulation of the Zeta VOIs (P.2) and multiple introductions of the Gamma VOC (P.1) in the municipality. We also highlight the incidence of cases in the community, identifying the most possible regions of the municipality and correlating clinical information with local variants. Our data suggest that symptoms related to more severe disease, such as dyspnea and prolonged fever, were approximately twice as common among Gamma/VOC and Zeta/VOI compared to the ancestral strains. Ageusia and anosmia were less common in the Gamma lineage and neurological manifestations were more common in the Zeta/VOI. The fatality rate and hospitalization rate were higher during the period of Gamma/VOC circulation. Our data suggest that Gamma/VOC and Zeta/VOI caused more severe disease and this may have contributed to the higher level of hospitalization and deaths during the second wave in Brazil. Furthermore, we

highlight how genomic epidemiology studies coupled with representative rate and detailed epidemiological data are key to providing solid insights into COVID-19 transmission.

Descriptors: SARS-CoV-2, genomic epidemiology, primary care, infectious diseases, COVID-19.

LISTA DE GRÁFICOS

Gráfico 1 . Evolução das variantes em São Caetano do Sul, São Paulo	67
Gráfico 2. Correlação do banco de dados.	71
Gráfico 3. Número efetivo de reprodução (Rt) por linhagem.	74
Gráfico 4. Análises descritivas de taxa de hospitalização, taxa de letalidade intra-hospitalar e taxa de hospitalização.	76
Gráfico 5 Probabilidade do sintoma ocorrer de acordo com a linhagem atribuída à amostra utilizando todos os casos sequenciados.	77
Gráfico 6. Probabilidade do sintoma ocorrer de acordo com a linhagem atribuída à amostra utilizando todos os positivos para COVID-19.	78
Gráfico 7. Probabilidade de ocorrência de determinado sintoma de acordo com a faixa etária e fase do estudo (I, II e III).	80

LISTA DE TABELAS

Tabela 1. Características demográficas e clínicas.	65
Tabela 2. Principais grupos de linhagens de SARS-CoV identificadas em São Caetano do Sul, agrupados por fase (I, II, III).	66

LISTA DE FIGURAS

Figura 1. A estrutura genômica e a árvore filogenética dos coronavírus.	22
Figura 2. Ciclo de vida.	24
Figura 3. Características da partícula de SARS-CoV-2.	33
Figura 4. Linhagens de SARS-CoV-2 ao longo do tempo.	38
Figura 5. Diagrama que representa as principais linhagens circulantes no Brasil.	41
Figura 6. Parâmetros epidemiológicos na transmissão de SARS-CoV-2.	46
Figura 7. O fluxograma contém informações sobre o desenho do estudo.	63
Figura 8. Análise filogenética das amostras de SARS-CoV-2 da cidade de São Caetano do Sul identificadas por bairro (A) e linhagem (B).	69
Figura 9. Inferência filogenética.	70
Figura 10. Mapas de incidência mensal e cumulativos em São Caetano do Sul.	72

LISTA DE SIGLAS E ABREVIATURAS

CoV - Coronavírus.

MERS-CoV - Síndrome Respiratória Aguda Grave do Oriente Médio.

SARS-CoV-1- Síndrome Respiratória Aguda Grave 1.

SARS-CoV-2 - Síndrome Respiratória Aguda Grave 2.

HCoV-OC43 - Coronavírus Humano OC43.

HCoV-HKU1 – Human coronavirus HKU1, originário de ratos.

HCoV-229E - Coronavírus humano 229E.

HCoV-NL63 - Coronavírus humano NL63.

Vírus (+) ssRNA vírus de RNA de cadeia simples e senso positivo.

TGEV - Transmissible Gastroenteritis Coronavírus.

BtCoV - Coronavírus de Morcego.

IBV - Vírus da Bronquite Infecciosa.

BCoV - Coronavírus bovino.

FCoV - Coronavírus felino.

CCoV - Coronavírus Canino.

RaTG13 - Coronavirus de morcego.

RmYN02 - Coronavírus de morcego.

ACE2 - receptor da enzima conversora de angiotensina 2.

ORF1 - do inglês, open reading frame 1, fase de leitura aberta 1.

ORF2 - do inglês, open reading frame 2, fase de leitura aberta 2.

RNA - ácido ribonucleico; em inglês: ribonucleic acid.

RdRP - RNA-dependent RNA polymerase.

mRNAs - RNA mensageiro.

sgRNAs - RNAs subgenômicos.

DMVs - vesículas de membrana dupla.

CMs - membranas convoluta.

DMSs: pequenas esferas de membrana dupla.

ER - Retículo Endoplasmático.

NSPs - proteínas não-estruturais.

RTC - complexo de transcrição-replicação.

ERGIG - compartimento intermediário Retículo Endoplasmático - Golgi.

OMS - Organização Mundial da Saúde.

PHEIC - Emergência de Saúde Pública de Interesse Internacional.

COVID-19 - Doença causada pelo coronavírus SARS-CoV-2.

SRAG - Síndrome Respiratória Aguda Grave.

RT-PCR - Transcrição reversa seguida de reação em cadeia da polimerase.

CDC - Center for Disease Control and Prevention.

LACENs - Laboratórios Estaduais Centrais de Saúde Pública.

HIV - Human immunodeficiency virus.

CT - Cycle threshold.

VOI - Variantes de Interesse.

VOC - Variantes de Preocupação.

VUM - Variante sob Monitoramento.

GISAID - Global Initiative for Sharing Avian Influenza Data.

NPIs - intervenções não farmacêuticas.

CHIKV - Chikungunya Vírus.

FAV - Febre Amarela Vírus.

EPI - Equipamentos de proteção individual.

R0 - Número de reprodução básico.

Rt ou Re - Número efetivo de reprodução.

EBOV - Vírus Ebola.

ECSA - East-Central-South-African.

IBGE - Instituto Brasileiro de Geografia e Estatística.

USCS - Universidade Municipal de São Caetano do Sul.

USP - Universidade de São Paulo.

UBS - Unidade Básica de Saúde.

UPA - Unidade de Pronto Atendimento.

API - Application Programming Interface.

SIVEP Gripe - Sistema de Informação de Vigilância Epidemiológica da Influenza

SH-aLRT - Shimoidara-Hasegawa.

MCMC - Markov Chain Monte Carlo.

SUMÁRIO

1. FUNDAMENTAÇÃO TEÓRICA	22
1.1 Histórico e diversidade dos coronavírus	22
1.2 Ciclo de vida	24
1.3 Transmissão de coronavírus entre espécies	27
1.4 Doenças humanas causadas por coronavírus	28
1.5 SARS (Síndrome Respiratória Aguda Grave)	29
1.6 MERS-CoV	30
1.7 SARS-CoV-2	31
1.8 Características genômica e evolutivas do SARS-CoV-2	32
1.9 Características clínicas	33
1.10 Diagnóstico	35
1.11 Perfil mutacional	35
1.12 Surgimento das variantes	37
1.13 Vacinologia	42
1.14 SARS-CoV-2 no Brasil	43
1.15 Epidemiologia de doenças infecciosas	45
1.16 Epidemiologia genômica	47
1.17 Filogenia, filodinâmica e filogeografia	48
2. OBJETIVOS	51
2.1 OBJETIVO GERAL	51
2.2 OBJETIVO ESPECÍFICO	51
2.3 MOTIVAÇÃO	51
3. MATERIAIS E MÉTODOS	52
3.1 Características do município de São Caetano do Sul (SCS)	52
3.2 Programa de atenção Primária Corona São Caetano	52
3.3 Seleção representativa de amostras para sequenciamento	53
3.4 Aspectos éticos	53
3.5 Diagnóstico de SARS-CoV	53
3.6 Amplificação viral	54
3.7 Sequenciamento de genoma	54
3.8 Montagem do genoma	55
3.9 Controle de qualidade de sequências de consenso do genoma	55
3.11 Análise filogenética e sinal temporal	56
3.12 Reconstruções filogeográficas bayesianas	57
3.14 Dados epidemiológicos e clínicos	57
3.15 Inferência da prevalência da linhagem ao longo do tempo	58
3.16 Estimativa do número de reprodução efetiva	59

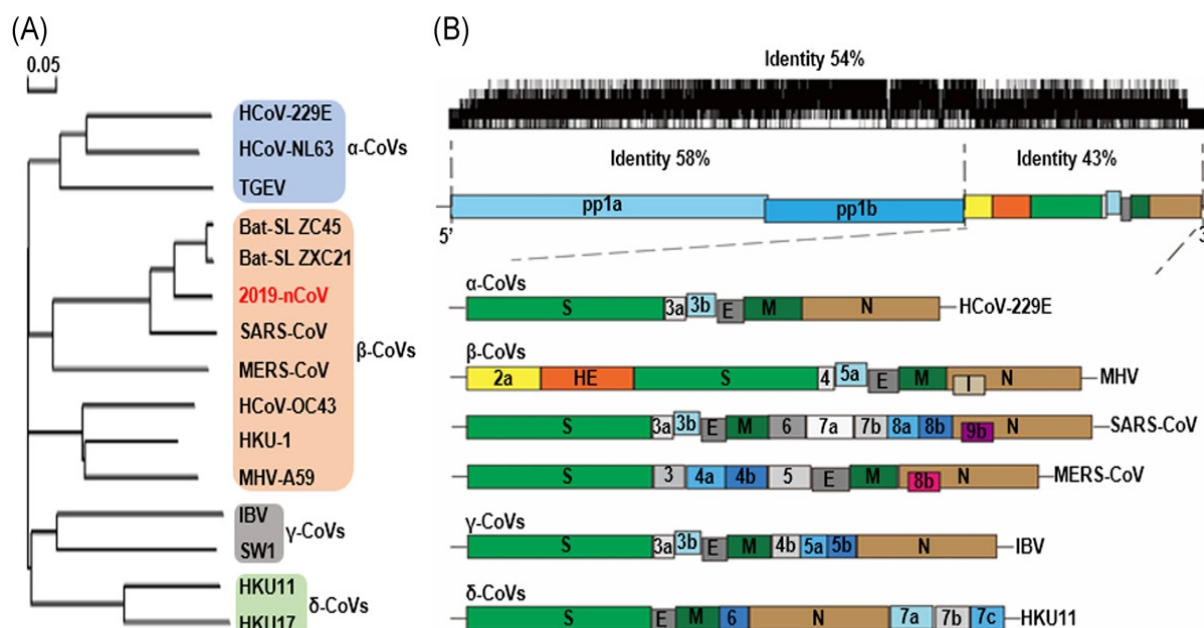
3.17 Estimativa de probabilidades de sintomas	60
4. RESULTADOS	63
4.1 Contexto epidemiológico	63
4.2 Análise filogenética	68
4.3 Inferências filogeográficas de relógio molecular	70
4.4 Análises epidemiológicas	72
5. DISCUSSÃO	81
6. CONCLUSÃO	85
7. MATERIAL SUPLEMENTAR	86
9. ANEXO I	97
10. ANEXO II	101
11. REFERÊNCIAS BIBLIOGRÁFICAS	109

1. FUNDAMENTAÇÃO TEÓRICA

1.1 Histórico e diversidade dos coronavírus

Os coronavírus (COVs) são membros da ordem *Nidovirales*, da família *Coronaviridae* e da sub-família *Coronavirinae* (1). Os Coronavírus são os mais importantes da ordem *Nidovirales* em doença humana e essa subfamília consiste em quatro gêneros - Alphacoronavirus, Betacoronavirus, Gammacoronavirus e Deltacoronavirus - com base em suas relações filogenéticas e estruturas genômicas. Enquanto os alfa (inclui HCoV-NL63 e HCoV-229E) e beta-coronavírus (inclui MERS-CoV, SARS-CoV-1, SARS-CoV-2, HCoV-OC43 e HCoV-HKU1) se originam em mamíferos, em particular, os morcegos; os gamma-coronavírus infectam principalmente aves e os delta-coronavírus além de infectar aves, alguns deles também podem infectar mamíferos, como os porcos¹ (Figura 1) (2).

Figura 1 A estrutura genômica e a árvore filogenética dos coronavírus



¹ Delta-coronavirus são extremamente importantes na pecuária.

Legenda: (A) A árvore filogenética de CoVs representativos, com o novo coronavírus 2019-nCoV destacado em vermelho. **(B)** A estrutura do genoma de quatro gêneros de coronavírus. Pp1a e pp1b representam os dois polipeptídeos longos que são processados em 16 proteínas não estruturais. As siglas S (*Spike*)², E (Envelope), M (Membrana) e N (Nucleocapsídeo) indicam quatro proteínas estruturais. A sigla HE (hemaglutinina-esterase) (2).

Todos os vírus da ordem *Nidovirales* contêm genomas muito grandes em comparação aos demais vírus de RNA, sendo que os vírus da família *Coronaviridae* são os maiores vírus de RNA descritos até o momento. Possuem cerca de 100nm de diâmetro, com genomas poliadenilados³ e capeados⁴ variando entre 25-32 kb. Eles são envelopados⁵, não segmentados⁶ de fita simples, Os vírus de RNA de sentido positivo⁷ (ssRNA +) e seus *vírions* são esféricos com uma grande glicoproteína, Spike (S), que se estende do envelope, resultando em sua forma de coroa (3). A análise de sequência mostra que o 2019-nCoV possui uma estrutura genômica típica de CoV e pertence ao clado de betacoronavírus que inclui Bat-SARS-like (SL)-ZC45, Bat-SL ZXC21, SARS-CoV e MERS-CoV. Com base na árvore filogenética de CoVs, 2019-nCoV está mais intimamente relacionado ao bat-SL-CoV-ZC45 e bat-SL-CoV-ZXC21 e mais distante ao SARS-CoV, conforme demonstrado na Figura 1A (2).

Os coronavírus são particularmente importantes e conhecidos por causar doenças em humanos e animais, causando cerca de um terço dos “resfriados comuns” e sendo reconhecidos pelas síndromes respiratórias agudas graves como o SARS, MERS e atualmente, o COVID-19. Mesmo sendo diagnosticados pela primeira vez há mais de 60 anos, os coronavírus só receberam notoriedade em 2003, quando um de seus membros foi identificado como o agente etiológico da Síndrome Respiratória Aguda Grave (SARS). Anteriormente, esses vírus eram conhecidos por serem importantes agentes de

² Spike em português, espícula.

³ Possui cauda Poli-A na extremidade 3'.

⁴ Possui um nucleotídeo de guanina metilado ligado ao RNAm via uma ligação trifosfato não usual de 5' para 5'

⁵ Vírus recobertos por uma camada de lipídios (gorduras).

⁶ Genoma composto por apenas um fragmento de RNA

⁷ Mesmo sentido do RNAm (RNA mensageiro)

infecções respiratórias e entéricas em animais domésticos, tais como, o TGEV (Transmissible Gastroenteritis Coronavírus, em português, Coronavírus da Gastroenterite Transmissível), que causa a gastroenterite suína; o Vírus da Bronquite Infecciosa (IBV), que causa doenças renais e do trato respiratório superior em galinhas e o coronavírus bovino (BCoV), que causa doença do trato respiratório e diarreia em bovinos, são todos patógenos economicamente importantes. Nos felinos e gatos domésticos, o coronavírus felino (FCoV) é altamente virulento (Sharun et al. 2020; Perlman and Netland 2009; Rottier et al. 2005).

1.2 Ciclo de vida

Por serem parasitas intracelulares obrigatórios, todos os vírus, incluindo os coronavírus, dependem da maquinaria de tradução da célula hospedeira para a produção de suas proteínas e progênie⁸ infecciosa. O ciclo de vida dos CoVs envolve quatro estágios principais: (i) ligação e entrada, (ii) expressão da replicase, (iii) replicação e transcrição e (iv) montagem e liberação e são descritos abaixo (**Figura 2**)

Figura 2 : Ciclo de vida

⁸ descendentes

sg) através das DMVs⁹, CMs¹⁰, DMSs¹¹. As proteínas estruturais migram das membranas do Retículo Endoplasmático (ER) para o Complexo de Golgi. Finalmente, os *vírions* são secretados da célula infectada por exocitose. As caixas em vermelho representam alvos antivirais (4).

1. A proteína S (*Spike*) do vírus se liga ao receptor (enzima conversora de angiotensina 2, no SARS-Cov-2) na superfície de uma célula epitelial. A interação proteína S-receptor é o fator determinante para um coronavírus infectar uma espécie hospedeira e também governa o tropismo tecidual do vírus¹². Esse receptor pode variar entre os diferentes coronavírus (ver tópico anterior).

2. Após a ligação do receptor, o vírus tem acesso ao citoplasma da célula hospedeira e o vírus entra no endossomo (endocitose).

3. Bombeamento dos Íons hidrogênio e diminuição do pH desses endossomos mudam a conformação da proteína S formando a subunidade S2 e permitindo a fusão entre a membrana viral com a membrana endossômica.

4. O genoma viral é de sentido positivo (mesmo sentido de um RNAm) e tem as características de um RNAm (Ele é capeado pela 5' metil guanosina e tem uma cauda 3' poli-A). Ribossomos se ligam ao RNA genômico que é usado como molde para traduzir duas ORFs 1a e 1b. Essas poliproteínas p1a e p1b são clivadas através de proteases¹³ para formar de 15 a 16 proteínas não-estruturais (NSPs). Algumas NSPs se associam para formar a RNA polimerase viral (replicase) através de um complexo de transcrição-replicação (RTC) nas vesículas de membrana dupla (DMVs) para criar um ambiente adequado para a replicação do RNA e transcrição dos RNAs subgenômicos (sg RNA)(4,5).

5. O RTC direciona a produção de RNAs de sentido negativo por meio de replicação e transcrição, que servem como molde para a geração de genomas de RNAs de sentido positivo. Durante a transcrição, um conjunto de 7 a 9 RNAs subgenômicos (sgRNAs), incluindo aqueles que codificam todas as

⁹ DMVs vesículas de membrana dupla

¹⁰ CMs membranas convolutas

¹¹ DMSs pequenas esferas de membrana dupla

¹² significa a propensão que um vírus tem em infectar determinado tipo de célula ou tecido em especial.

¹³ Enzimas que quebram ligações peptídicas entre os aminoácidos das proteínas

proteínas estruturais (S, E, M e N) e proteínas acessórias, são produzidos por transcrição descontínua¹⁴. Esta natureza descontínua tem como consequência um alto grau de recombinação resultante da inserção de sequências virais e não virais ou frequentes deleções no genoma, o que pode levar tanto à formação de genomas viáveis como de RNAs interferentes defeituosos (4,6).

6. Enquanto o processo de replicação está em andamento, os vírus progênes começam a se reunir e as proteínas estruturais são traduzidas e migram para o retículo endoplasmático e subsequentemente, para as membranas do Retículo Endoplasmático - Golgi (ERGIC). A proteína estrutural N encapsula esses genomas recém produzidos e torna os *virions* maduros. As proteínas M e E são responsáveis pela produção de envelopes virais. Adicionalmente, as proteínas M, S, juntamente com o nucleocapsídeo, promovem a conclusão da montagem do vírion. Finalmente, os vírions são secretados da célula infectada por exocitose¹⁵ (4,5).

1.3 Transmissão de coronavírus entre espécies

Uma característica importante da família dos coronavírus é a capacidade de cruzar as barreiras entre espécies (7), fenômeno conhecido como *spillover* - em português, transbordamento. O coronavírus humano OC43 (HCoV-OC43) e o coronavírus bovino (BCoV) estão intimamente relacionados e acredita-se que o vírus tenha se originado em uma espécie e depois em espécies cruzadas. O BCoV também se espalhou para vários outros animais, como alpaca e ruminantes selvagens. Um outro exemplo de transmissão entre espécies, seria o Coronavírus Felino I (FCoV-I) e o Coronavírus Canino I (CCoV-I) que compartilham um ancestral comum, assim como o CCoV-I, que recombinau-se com um outro coronavírus desconhecido e deu origem ao coronavírus canino II (CCoV-II). Para o surgimento do coronavírus felino II (FCoV-II), houve uma recombinação do CCoV-II e do FCoV-I (hospedeiro desconhecido). Especula-se também que o CCoV-II se espalhou para os porcos resultando no vírus da gastroenterite suína (TGEV) (7).

¹⁴ Processo único aos vírus de RNA que ocorre durante a síntese de RNAs de fita negativa

¹⁵ É um processo pelo qual a célula expõe material para fora do citoplasma através da membrana celular

Para as doenças humanas, o padrão de recombinação entre espécies também ocorre. No caso da epidemia de SARS-CoV, um Coronavírus de Morcego semelhante (BtCoV), se espalhou e adaptou-se em animais selvagens, como o civeta (*Paguma larvata*), que era vendido nos mercados chineses como alimento; observou-se que os tratadores desses animais selvagens se infectaram com o vírus, porém, o vírus não causava doença. Após adaptações, resultou-se em cepas que se replicaram com eficiência no hospedeiro humano e podiam se espalhar de pessoa para pessoa, causando doenças. A epidemia de SARS-CoV foi identificada quando um médico que atendia esses tratadores do mercado foi infectado e transmitiu para outros contactantes (8). Em relação ao MERS (Síndrome Respiratória do Oriente Médio), a transmissão ocorreu dos camelos dromedários para os humanos (9), provavelmente espalhou-se de morcegos para camelos dromedários há pelo menos 30 anos e, desde então, tem sido prevalente em camelos dromedários (10). Além do SARS-CoV, existem outros exemplos de transmissão cruzada de espécies de coronavírus, os vírus BCoV e HCoV-OC43 são semelhantes e o vírus pode ter passado de hospedeiros bovinos para humanos há aproximadamente 100 anos (11). Assim, os morcegos são provavelmente os principais reservatórios naturais dos alfa e beta-coronavírus (12). Com base nos bancos de dados de sequências atuais, todos os coronavírus humanos têm origens animais. (13,14). Outros estudos de vigilância genômica identificaram mais de 60 novos coronavírus de morcegos em vários pontos do mundo, como na África (15), Europa (16), América do Norte (17), China (18,19) e mais recentemente, na Rússia (20). Além dos coronavírus encontrados em morcegos, grupos de pesquisa identificaram novos coronavírus em outros animais como, leopardos asiáticos (21) e baleia de cativeiro (22).

1.4 Doenças humanas causadas por coronavírus

Antes da epidemia de SARS em 2002-2003, somente dois outros coronavírus eram conhecidos por importantes infecções humanas: HCoV-OC43 e HCoV-229E (Coronavírus humano 229E) que infectam o trato respiratório superior em hospedeiros imunocompetentes, idosos ou recém-nascidos e causam sintomas relativamente menores (23). Após a epidemia de SARS-CoV,

intensificou-se a busca por novos coronavírus relacionados à doenças humanas, o que levou à identificação do HCoV-NL63 (Coronavírus humano NL63) (24) (25) e do HCoV-HKU1 (26). Esses quatro coronavírus humanos 229E, NL63, OC43 e HKU1, normalmente infectam apenas o trato respiratório superior em hospedeiros imunocompetentes e causam sintomas relativamente mais leves (27) Segundo Pyrc et al. (28), o NL63 infectou populações humanas por séculos, onde através de análises filogenéticas, ele se divergiu do 229E há quase 1.000 anos.

1.5 SARS (Síndrome Respiratória Aguda Grave)

No final de 2002, os primeiros casos de uma nova infecção do trato respiratório superior evoluindo rapidamente para pneumonia grave, foram relatados na província de Guangdong, na China e seu agente causador foi posteriormente descrito como o primeiro coronavírus humano patogênico (SARS-CoV). Acelerada pelas viagens aéreas e aglomerações em hospitais e regiões urbanas, a doença se espalhou rapidamente para várias partes do mundo. Estima-se que o SARS-CoV tenha causado mais de 8.000 casos e 774 mortes em pelo menos 26 países, figurando como a primeira pandemia do século 21. É um agente altamente infeccioso e tem um índice de fatalidade de 10%. Seus epicentros foram na China e Hong Kong, mas foi contida em 2003, embora alguns casos adicionais ocorreram em 2004 (29). De origem zoonótica, os morcegos são os prováveis reservatórios do vírus, enquanto as civetas de palma mascarada (*Paguma larvata*) e cães-guaxinim (*Nyctereutes procyonoides*) serviram como hospedeiros intermediários (30). O SARS-CoV tem um genoma de 29727 pares de bases e a organização do genoma é muito similar às de outros coronavírus (5' replicase (rep), espícula (S), envelope (E), membrana (M), nucleocapsídeo (N)-3' e pequenas regiões não traduzidas em ambos os terminais) (29). A transmissão ocorre por contato da mucosa com gotículas respiratórias ou objetos de indivíduos infectados e infecta principalmente células epiteliais do pulmão sendo capaz de entrar em macrófagos e células dendríticas (31). Os casos notificados ocorreram principalmente em ambientes hospitalares devido ao contato próximo aos pacientes (32), com surtos, às vezes, envolvendo mais de 100 indivíduos

(33,34). Os casos também foram associados a eventos de super disseminação (35). É uma doença caracterizada por uma febre de mais de 38 graus, acompanhada de dor de cabeça, mal estar geral e dores e que pode levar a pneumonia. Os sintomas respiratórios normalmente são amenos, mas o paciente pode desenvolver uma tosse não produtiva e juntamente com dispnéia¹⁶ (29).

1.6 MERS-CoV

Quase uma década após o surto controlado de SARS-CoV, emergiu o próximo coronavírus zoonótico: Síndrome Respiratória do Oriente Médio (MERS-CoV)¹⁷. Foi descrito em Junho de 2012 na Arábia Saudita, em um paciente que apresentou uma pneumonia incomum, seguido de falência múltipla de órgãos e morte (36). O MERS-CoV é um vírus zoonótico, estudos mostraram que os humanos são infectados através do contato direto ou indireto com camelos dromedários infectados. A transmissão de humano para humano é possível e ocorreu predominantemente entre contatos próximos, em ambientes de assistência à saúde. Fora do ambiente hospitalar, não houve transmissão sustentada de humano para humano documentada em qualquer lugar do mundo. Outro fator importante, é que os casos identificados fora do Oriente Médio são geralmente indivíduos que parecem ter sido infectados no Oriente Médio e depois viajaram para áreas fora da região. Até o momento, um número limitado de surtos ocorreu fora do Oriente Médio (37). No final de 2021, um total de 2578 casos de MERS e 888 mortes foram relatados em todo o mundo, com quase 85% dos casos confirmados na região da Arabia Saudita e um índice de fatalidade muito alto em humanos, de cerca de 35%. Os sintomas típicos incluem febre, tosse e falta de ar, alguns pacientes apresentaram pneumonia. Sintomas gastrintestinais, incluindo diarreia, também foram relatados nos pacientes. Desde 2012, um total de 15 países relataram casos confirmados de MERS na região do Mediterraneo e 27 relataram casos. (38) (37). Os países mais afetados implementaram medidas de controle importantes que resultaram na diminuição no número de casos localmente e

¹⁶ falta de ar, dificuldade em respirar

¹⁷ Do inglês, Middle East Respiratory Syndrome Coronavirus

conseqüentemente, globalmente (39). Essas epidemias internacionais de SARS-CoV e MERS-CoV encorajaram a comunidade científica internacional a investigar mais coronavírus altamente patogênicos. Desde então, vários novos coronavírus foram descritos em morcegos e em outros animais, sugerindo que o transbordamento de novos coronavírus para populações humanas é muito provável de ocorrer.

1.7 SARS-CoV-2

Casos de pneumonia grave de etiologia desconhecida em Wuhan, província de Hubei, foram relatados em 31 de dezembro de 2019 pelo Escritório da OMS (Organização Mundial da Saúde) na China. Esses casos estavam inicialmente ligados ao Mercado de Frutos do Mar de Huanan, onde a maioria dos pacientes trabalhavam ou tinham visitado recentemente. Até 4 de janeiro de 2020, um total de 44 pacientes foram relatados apresentando febre. No entanto, 11 pacientes apresentaram sintomas graves como dificuldade em respirar e lesões pulmonares invasivas (40). Alguns dias depois, cientistas chineses relataram ter identificado um novo coronavírus, nomeado como 2019-nCoV, suspeito de ser o agente causador dos casos relatados em Wuhan (41,42). A primeira sequência genômica do novo coronavírus foi compartilhada pelos cientistas chineses em 10 de janeiro de 2020 (43) e no dia seguinte, foi relatada a primeira morte associada à nova pneumonia, um homem de 61 anos que frequentava o mercado de Frutos do Mar de Huanan, Wuhan (41,42). Enquanto isso, os casos se espalharam para outras províncias chinesas e em 13 de Janeiro de 2020, autoridades da Tailândia relataram o primeiro caso de novo coronavírus de um viajante que acabava de retornar de Wuhan (44). Logo em seguida, outros casos importados foram relatados no Japão (45), Estados Unidos (46) e na França (47). Todos os pacientes tinham histórico de viagem recente à China.

Devido ao aumento exponencial de casos relatados, o governo chinês isolou a cidade de Wuhan e alguns dias depois, toda a província de Hubei, uma manobra para evitar uma maior disseminação nacional e internacional de casos de COVID-19 (48). Considerando o aumento de casos, a OMS emitiu uma declaração de Emergência de Saúde Pública de Interesse Internacional

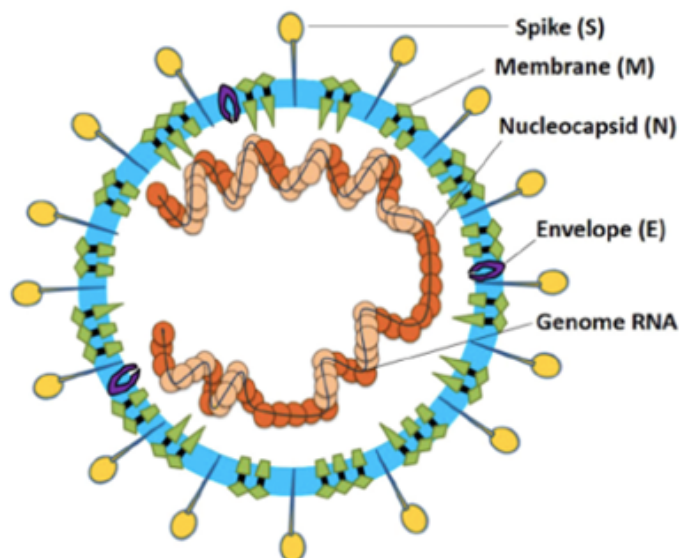
(PHEIC) no final de Janeiro de 2020, porém, nesse momento já haviam mais de 7800 casos confirmados em todo o mundo (49). Na primeira quinzena de Março de 2020, a OMS emitiu declaração de pandemia e o mundo atingiu mais de 100 mil casos confirmados da doença, com a Europa passando a ser o epicentro de casos de COVID-19, sendo a Itália e a Espanha os países mais atingidos (50).

1.8 Características genômica e evolutivas do SARS-CoV-2

O SARS-CoV-2 foi identificado e sequenciado no início de 2020 (41,42,51). O SARS-CoV-2 possui um genoma com aproximadamente 30.000 pb (Figura 5) e apresenta uma similaridade de 89,1% com um outro coronavírus do tipo SARS (gênero *Betacoronavirus*, subgênero *Sarbecovirus*) que já havia sido encontrado em morcegos na China (47,48). A cepa de referência do SARS-CoV-2, Wuhan-Hu-1 (acesso ao GenBank MN908947), foi identificada em um paciente em Wuhan, na China, em 26 de dezembro de 2019. Assim, como outros coronavírus, o SARS-CoV-2 codifica várias proteínas estruturais e não estruturais. São quatro proteínas estruturais, conhecidas como proteínas S (*spike*), E (envelope), M (membrana) e N (nucleocapsídeo). Na superfície das partículas do vírus é expressa a proteína S (Spike) em forma de espículas, que tem uma aparência característica de “coroa” (do latim: corona = coroa), ela é responsável pela infecção da célula hospedeira mediando a ligação e a fusão da membrana celular por meio do receptor da enzima conversora de angiotensina 2 (ACE2). A proteína S compreende duas subunidades: S1, responsável pela ligação ao receptor e, um domínio S2, responsável pela fusão da membrana celular (51)

O SARS-CoV-2 é um vírus RNA envelopado de fita simples positiva que evolui a uma taxa inferior a 1×10^{-3} substituições/local/ano, totalizando ~ 2 substituições por genoma por mês, uma taxa semelhante à taxa evolutiva média observada em outros genomas de vírus de RNA (52). A taxa é mais lenta do que a maioria dos vírus de RNA devido à atividade de revisão codificada pelo gene não-estrutural NSP14, portanto, o SARS-CoV-2 pode não sofrer mutações tão rápido quanto os vírus da gripe A, por exemplo, mas ainda assim, possui populações heterogêneas e dinâmicas (53).

Figura 3. Características da partícula de SARS-CoV-2.



Legenda. Estrutura do vírus SARS-CoV-2. Os coronavírus são genomas de vírus de RNA de fita simples, envelopados, não segmentados e de sentido positivo, com tamanhos que variam de 26 a 32 quilobases. O vírion possui um nucleocapsídeo composto de RNA genômico e proteína de nucleocapsídeo fosforilado (N), uma glicoproteína spike (S). A proteína da membrana (M) (uma glicoproteína transmembrana tipo III) e a proteína do envelope (E) estão localizadas entre as proteínas S no envelope do vírus (54).

1.9 Características clínicas

Os humanos não têm experiência imunológica direta com o SARS-CoV-2, o que nos deixa vulneráveis a infecções e doenças. Assim como os outros coronavírus respiratórios, o SARS-CoV-2 é transmitido principalmente por gotículas respiratórias., porém, as partículas virais também estão presentes em aerossóis, sangue, secreções oculares, urina e fezes, mas se espalham principalmente por transmissão respiratória direta de pessoa para pessoa. Indivíduos acometidos pela COVID-19 manifestaram uma vasta gama de sintomas, variando de sintomas leves a doenças graves. Os sintomas podem aparecer entre 2 a 14 dias após a exposição ao vírus (55). Estima-se que a maioria dos pacientes com COVID-19 sejam assintomáticos ou apresentem

quadro leve. Os sintomas mais comuns relatados são: febre, tosse seca, dificuldade em respirar, dores musculares e/ou articulares, dor de cabeça, dor de garganta, tontura, diarreia, náusea, fadiga. Anosmia¹⁸ e ageusia¹⁹ também podem ser observadas (56,57). A febre pode estar presente em aproximadamente metade dos pacientes no momento da apresentação inicial, dependendo do estudo, mas, em geral, a literatura sugere que 20 a 99% dos pacientes apresentam febre durante o curso da doença (55,58%). Os sintomas podem mudar com novas variantes do COVID-19 e podem variar dependendo do estado de vacinação. Pacientes infectados com a linhagem Omicron, por exemplo, apresentaram mais tosse, coriza/nariz entupido e fadiga/letargia (59%). Já pacientes infectadas com a linhagem Delta estão associadas a piores desfechos, incluindo necessidade de hospitalização, internação em UTI e mortalidade (55).

A gravidade da doença nos pacientes deve-se não apenas à infecção viral, mas também à resposta do hospedeiro. O padrão de gravidade da doença é acentuado pela idade e pelos fatores de risco que incluem: idade superior a 75 anos, diabetes, câncer, história de transplante, hipertensão e doença cardíaca ou pulmonar prévia (55,60). A Obesidade está ligada a maior necessidade de internação (68,69). Os casos graves de COVID-19 evoluem para a Síndrome Respiratória Aguda Grave (SRAG), em média cerca de 8 a 9 dias após o início dos sintomas. A SRAG observada em COVID-19 é caracterizada pela dificuldade em respirar e baixo nível de oxigênio no sangue que pode levar diretamente à insuficiência respiratória, sendo a causa da morte em 70% dos casos fatais de COVID-19 (61). Outros fatores podem contribuir para a piora do quadro clínico do paciente e aumentam a possibilidade de hospitalização. A possibilidade de coinfeção com outros vírus também foi observado (60). Os vírus respiratórios mais frequentemente identificados em pacientes com diagnóstico suspeito de COVID-19 foram os vírus da Influenza A/B e Rinovírus Humanos (62) (63). Co-infecções bacterianas ou fúngicas são uma importante fonte de morbidade e mortalidade que podem afetar até 8% dos pacientes (64).

¹⁸ anosmia: perda de olfato.

¹⁹ Ageusia: perda de paladar.

1.10 Diagnóstico

O RT-PCR em tempo real (rRT-PCR) é usado rotineiramente na detecção qualitativa de ácido nucleico do SARS-CoV-2 para testes de diagnóstico COVID-19 em amostras de swabs nasofaríngeos, orofaríngeos e saliva através de protocolos recomendados pela OMS ou CDC (Center for Disease Control and Prevention) (65,66). Os *primers*²⁰ usados no rRT-PCR são fundamentais para o diagnóstico preciso do COVID-19 e a descoberta de novas linhagens. As sequências de *primers* são especialmente desenhadas para amplificar as regiões conservadas (o gene da RNA polimerase dependente de RNA (RdRP) na região ORF1ab; o gene da proteína E e o gene da proteína N) nas diferentes linhagens existentes para alta especificidade e sensibilidade e também estão sujeitas a alterações de genótipo à medida que o coronavírus SARS-CoV-2 evolui, necessitando de atualização de acordo com as regiões conservadas e as mutações predominantes nos genomas do SARS-CoV-2 (67). Ao longo da pandemia, testes rápidos baseados em antígeno que detectam proteínas virais, ficaram disponíveis para a população, o que ajudou a expandir a capacidade de teste para ambientes fora do hospital ou da clínica (68). Nesse processo, a coleta da amostra também é realizada via swab oro/nasofaríngeo, porém, o resultado sai em poucos minutos, dispensando grandes infra-estruturas. Apesar da especificidade geralmente alta, o teste SARS-CoV-2-Ag demonstra sensibilidade altamente variável, provavelmente devido a cargas virais variáveis (69).

1.11 Perfil mutacional

Os vírus de RNA são os organismos biológicos bem-sucedidos devido a capacidade de mutação²¹ que permitem o surgimento de novas doenças, sua rápida disseminação e um alto potencial adaptativo que incluem ciclos rápidos de replicação, enorme tamanho populacional e extensa diversidade genética. Vírus como HIV e Influenza podem também demonstrar resistência a antivirais

²⁰ *primers*: iniciadores são segmentos de ácidos nucleicos necessários à iniciação da replicação do material genético.

²¹ alteração de aminoácidos

e vacinas. Por conta desse alto potencial adaptativo, a maioria dos genomas de vírus de RNA não excedem 15kb, com exceção dos coronavírus que codificam genomas de até 32kb. A evolução dos vírus de RNA é limitada para garantir estabilidade do genoma, pois a maioria das mutações aleatórias são deletérias e podem resultar em menor aptidão. Grandes vírus de RNA, como os pertencentes à família Coronavírus, parecem ter um sistema de revisão que melhoram a fidelidade da polimerase²² e reduzem as taxas de mutações (70). Os coronavírus contêm uma enzima exonuclease que reduz sua taxa de erro de replicação em cerca de 15 a 20 vezes *in vitro*, resultando em uma taxa de mutação viral *in vivo*, cerca de 10 vezes menor do que a da Influenza (71). Apesar desse sistema de revisão, os coronavírus também acumulam mutações e geram mais diversidade através do processo de recombinação, quando variantes com mutações diferentes infectam o mesmo hospedeiro (72). O vírus SARS-CoV-2 tem uma taxa evolutiva estimada que varia de 0,8 a 2×10^{-3} substituições por local por ano (s/s/y) (73,74). Isso equivale a aproximadamente a uma mutação a cada duas semanas ou 2-3 mutações por mês. Apesar do aumento da imunidade da população, o número crescente de variantes do SARS-CoV-2 compartilham mutações que permitem que o vírus se espalhe, mantendo ou aumentando sua capacidade de replicação (71). No entanto, a classificação filogenética de linhagens emergentes de SARS-CoV-2 tem sido difícil porque novas linhagens geralmente diferem umas das outras por apenas alguns nucleotídeos (75).

O SARS-CoV-2 foi observado em várias espécies de animais domésticos e selvagens (como resultado de eventos zoonóticos reversos), e essas novas variantes podem ter evoluído em um desses hospedeiros alternativos e voltou para a população humana ou essas variantes podem ter evoluído em indivíduos imunocomprometidos ou imunossuprimidos com infecções crônicas, como o HIV, por exemplo (76). O que reforça a importância do sequenciamento na questão de vigilância genômica para identificação de variantes preocupantes mais rapidamente e criar medidas que controlem a disseminação (77).

²² precisão com que a informação genética, seja DNA ou RNA, é copiada em relação a uma sequência referência

1.12 Surgimento das variantes

Inicialmente, identificou-se duas grandes linhagens de SARS-CoV-2, que circulavam desde o início do surto em Wuhan, sugerindo que pelo menos dois eventos de transbordamento independentes poderiam ter acontecido (linhagens A e B). As sequências da linhagem B diferem em duas posições de nucleotídeos (8.782 em ORF1ab e 28.144 em ORF8) das sequências mais próximas de vírus de morcegos conhecidos (RaTG13²³ e RmYN02²⁴), enquanto os vírus da linhagem A têm essas posições conservadas, sugerindo que a linhagem A é a mais próxima ancestral comum das cepas de SARS-CoV-2 circulantes humanas (71,78). Embora essas linhagens se espalharam globalmente, a linhagem B tornou-se dominante em todo o mundo. A primeira indicação de que o SARS-COV-2 estava se adaptando aos humanos foi a substituição em nível populacional do alelo 614D pelo alelo 614G (mutação de nucleotídeo A-para-G na posição 23.403 na cepa de referência Wuhan) na proteína Spike, no início de 2020 e que foi passada até as linhagens B subsequentes em várias localizações geográficas e que atingiu sua prevalência de quase 100% em junho de 2020 (77,79).

No final de 2020, também foram relatadas várias variantes com mutações recorrentes (além de D614G) ocorrendo principalmente, mas não exclusivamente, na proteína *Spike* (79). Outras mutações têm sido associadas ao surgimento de linhagens de potencial interesse em saúde pública que levou à caracterização dessas variantes. Tais linhagens são atualmente separados em três categorias diferentes de acordo com suas características epidemiológicas: Variantes de Preocupação (VOC), Variante de Interesse (VOI) e Variante sob Monitoramento (VUM), a fim de priorizar o monitoramento, pesquisa globais e informar a resposta em andamento à pandemia do COVID-19 com o objetivo de impedir sua propagação (80). Para identificação e vigilância epidemiológica das linhagens de SARS-CoV-2, dois sistemas foram criados, o sistema PANGO (78) e Nextclade (81). Esses sistemas fornecem uma visão geral das variantes e mutações do SARS-CoV-2 que são de

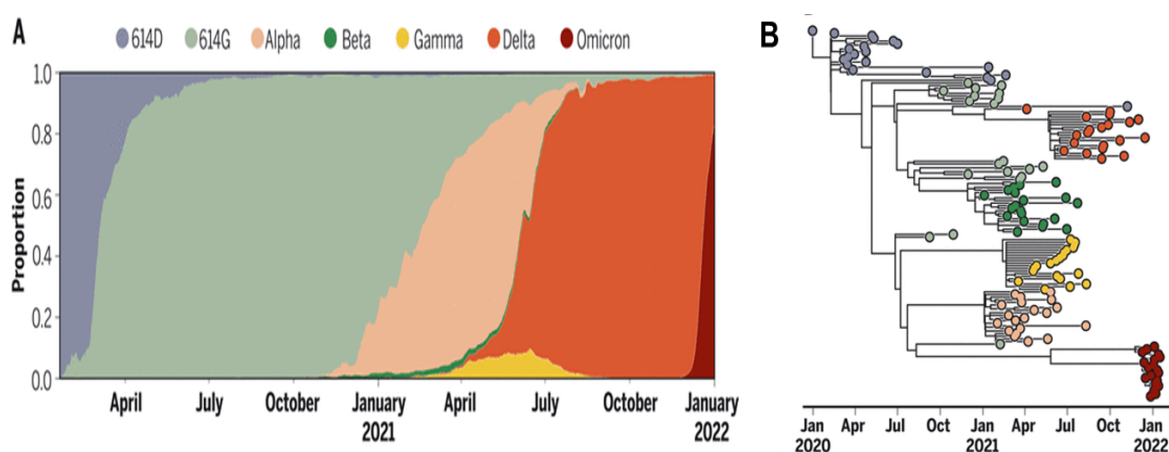
²³ um betacoronavírus que infecta o morcego ferradura, *rhinolophus affinis*, que possui 96% de similaridade com o SARS-CoV-2.

²⁴ um coronavírus de morcego que possui 93% de similaridade com o SARS-CoV-2.

interesse da comunidade científica. Os especialistas juntamente com a OMS elaboraram um sistema de nomenclatura dinâmica que consideraram rótulos fáceis de pronunciar e não estigmatizantes para VOI e VOC, recomendando o uso de letras do alfabeto grego (80).

Algumas características epidemiológicas são relevantes a fim de designar uma variante de SARS-CoV-2 como VOI que são previstas ou conhecidas por afetarem as características do vírus (como transmissibilidade, gravidade da doença, escape imunológico, escape diagnóstico ou terapêutico) e foi identificado por causar transmissão comunitária significativa e/ou múltiplos aglomerados de COVID-19, em vários países com prevalência relativa crescente juntamente com o aumento do número de casos ao longo do tempo, ou outros impactos epidemiológicos aparentes. Já a definição de VOC é apresentar as mesmas características de um VOI com a adição de que uma ou mais características alteradas que foram confirmadas por meio de estudos comparativos (transmissibilidade e virulência aumentadas, apresentação alterada da doença ou diminuição da eficácia das medidas de saúde pública, diagnósticos, terapêuticas ou vacinas). Por sua vez, os VUMs são variantes que abrigam mutações suspeitas que podem potencialmente representar um risco futuro mas carecem de evidências em relação ao seu impacto fenotípico ou epidemiológico exigindo monitoramento. Após investigação, pode ser considerado um VOI (80).

Figura 4. Linhagens de SARS-CoV-2 ao longo do tempo.



Legenda: A) As frequências de variantes do SARS-CoV-2 preocupantes ao longo do tempo. A primeira indicação de que o SARS-COV-2 estava se adaptando aos humanos foi a substituição em nível populacional do alelo 614D pelo alelo 614G no início de 2020. As frequências variantes no eixo y são calculadas com base nos dados da sequência do SARS-CoV-2 depositado no GISAID (Global Initiative for Sharing Avian Influenza Data). B) Apenas as principais linhagens variantes circulantes globalmente são mostradas. Uma filogenia inferida a partir de dados da sequência SARS-CoV-2 mostrando as relações evolutivas entre as linhagens variantes incluídas em (A)₍₇₇₎.

Um total de cinco VOCs foram identificadas até o momento: Alfa (B.1.1.7, GRY, 20I/V1), Beta (B.1.351, GH/501Y.V2, 20H/V2), Gama (P.1, GR/501Y.V3, 20J/V3), Delta (B.1.617.2, G/478K. V1/21A, 21I, 21J) e Omicron (B.1.1.529/GRA/21K, 21L, 21M). A primeira variante de preocupação identificada foi a linhagem Alfa (B.1.1.7) no sudeste da Inglaterra, descrita em 18 de dezembro de 2021, pelo grupo de vigilância COG-UK (COVID-19 Genomics UK Consortium) e se espalhou rapidamente pelo Reino Unido, apresentando um número extraordinariamente grande de alterações genéticas, particularmente na proteína Spike: são 14 mutações não sinônimas²⁵, 3 deleções²⁶ de nucleotídeos e 6 mutações sinônimas²⁷ ([Rambaut et. 2020](#)). Algumas características importantes dessa linhagem são: a mutação N501Y presente na região de domínio de ligação ao receptor (RBD) e a deleção nas posições 69 e 70 da proteína spike ($\Delta 69-70$) que foi associada à falha de testes de diagnóstico usando a sonda ThermoFisher TaqPath, que tem como alvo a proteína *spike* (82). As diversas mutações presentes nessa linhagem alteraram o fenótipo viral através do aumento da ligação à enzima conversora de angiotensina 2 (ACE2) (83) e consequentemente, o aumento da transmissibilidade (84) de 1,5 a 2 vezes em comparação com outras linhagens que estavam em circulação naquele momento (85). Dada a sua maior transmissibilidade, essa linhagem se espalhou para fora do Reino Unido e tornou-se a principal linhagem circulante, especialmente na Europa, América e

²⁵ não sinônimas: alteração em um par de bases de DNA que resulta na substituição de um aminoácido por outro na proteína produzida por um gene.

²⁶ Deleção: remoção de um ou mais nucleotídeos da sequência de DNA.

²⁷ Sinônimas ou silenciosas: mudança em um par de bases de DNA que altera o códon para outro que codifica o mesmo aminoácido e não causa mudanças na proteína produzida.

Oriente Médio. Os dois primeiros casos da América Latina, foram confirmadas em 31 de Dezembro de 2021 em São Paulo, através da vigilância genômica (86).

Após a descrição da linhagem Alfa pelo governo do Reino Unido, duas novas VOCs foram identificadas na África do Sul (Beta) e no Brasil (Gama), ambas apresentando uma constelação de mutações, incluindo N501Y. A linhagem Beta (B.1.351) também foi descrita pela primeira vez em dezembro de 2020 na África do Sul e detectada em amostras a partir de 8 de outubro de 2020. Essa linhagem é definida por oito mutações na proteína *spike*, incluindo três substituições (K417N, E484K e N501Y) e uma deleção (87). Foi a linhagem responsável pela segunda onda, ultrapassando os 20.000 casos diários e um aumento de cerca de 20% de mortalidade hospitalar, substituindo rapidamente as linhagens que circulavam anteriormente (B.1.1.54, B.1.1.56 e C.1). Outro fato interessante, é que esse aumento ocorreu em um cenário em que se estimava que mais de 30% da população já estava infectada (88).

No Brasil, tivemos nossa própria variante de preocupação identificada em Manaus, no Estado do Amazonas, em amostras coletadas entre 15 e 23 de dezembro de 2020 (89) e também identificada em viajantes que chegam a Tóquio - Japão, vindo do Amazonas (90). A linhagem Gama (P.1) carrega uma constelação distinta de mutações definidora de linhagens, onde possui 17 mutações, incluindo três na proteína *spike* (K417T, E484K e N501Y), associada ao aumento da ligação ao receptor humano ACE2 (enzima conversora de angiotensina 2). A análise filogenética indicou que P.1 - e outra linhagem, P.2 (91) - foram descendentes da linhagem B.1.1.28, detectada pela primeira vez no Brasil, no início de março de 2020. Essa nova linhagem é filogeneticamente e geneticamente distinta dos vírus ancestrais e carrega mutações que são associadas à uma rápida disseminação, evasão imune e neutralização de anticorpos (92). Devido ao surgimento da linhagem Gama, Manaus experimentou uma segunda onda de magnitude sem precedentes de casos e óbitos, espalhou-se para todos os outros estados brasileiros e foi a linhagem dominante em toda América Latina (93) no primeiro e segundo trimestre de 2021.

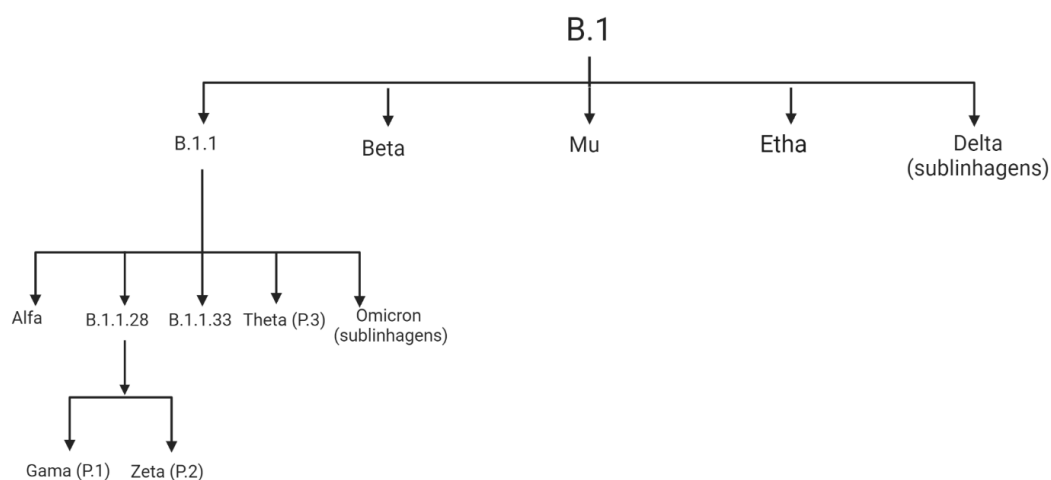
Apesar da população apresentar um alto nível de positividade imunológica, uma outra variante dominou o mundo no final de 2021, a variante

de preocupação Delta (B.1.617.2) e seus descendentes, identificada primeiramente na Índia (94). A variante Delta superou suas antecessoras (B.1.617.1/Kappa e B.1.1.7/Alfa) por mutações que aumentaram a replicação, a evasão imune, a avidéz do receptor do hospedeiro, aumentando assim a transmissibilidade, reinfecção e avanço da vacinação. A linhagem Delta é caracterizada pelas mutações T19R, Δ 157-158, L452R, T478K, D614G, P681R e D950N. Algumas dessas mutações demonstraram afetar a replicação e/ou transmissão do vírus, como P681R, que está localizada no local de clivagem das subunidades de pico S1-S2 e leva ao aumento da replicação e transmissão (95). Em estudos *in vitro*, essa linhagem demonstrou ser 6x menos sensível aos anticorpos neutralizantes do soro de indivíduos recuperados e 8x menos sensível aos anticorpos induzidos pela vacina, em comparação com a linhagem selvagem Wuhan-1 com a mutação D614G (96). Devido a essas características, rapidamente se tornou a linhagem dominante em todo mundo, mesmo em países onde outras VOCs já estavam estabelecidas (94,97).

A variante Omicron e suas sublinhagens (B.1.1.529) foram identificadas por pesquisadores da África do Sul, em novembro de 2021 (98). Possui um perfil mutacional que consiste em mais de 30 mutações na glicoproteína *Spike* em relação ao genótipo de referência Wuhan-Hu-1 (99). Os primeiros estudos confirmaram a capacidade da Omicron de escapar de anticorpos de infecção e vacinação anteriores e indicaram que ele pode usar uma via alternativa de entrada celular, o que melhora sua capacidade de infectar células no trato respiratório superior (100). Em dezembro de 2021, esta variante se espalhou rapidamente por todos os países, substituiu a variante Delta, resultando no maior pico de infecções confirmadas em diversos países e conseqüentemente, pico de mortes atribuídas ao vírus, principalmente em indivíduos não vacinados (77).

Figura 5. Diagrama que representa as principais linhagens circulantes no Brasil.

Principais linhagens identificadas no Brasil



Legenda. Diagrama que representa as principais linhagens circulantes no Brasil até 10 de Março de 2023. A linhagem ancestral principal é B.1. A linhagem B.1.1.28 teve um impacto crucial levando a variantes que apresentaram alta frequência no Brasil genomas (Gamma e Zeta). Adaptado de (81)

1.13 Vacinologia

Pela primeira vez, na história da vacinologia, numa velocidade surpreendente, as vacinas contra SARS-CoV-2 foram desenvolvidas, testadas e aprovadas para uso emergencial antes do tempo previsto. Com pouca imunidade adquirida ao SARS-CoV-2 e uma alta incidência de COVID-19 em meados de 2020, os fatores epidemiológicos facilitaram o cronograma acelerado desses ensaios cruciais que demonstraram uma ótima proteção contra o COVID-19 e contribuíram para diminuição dos casos graves, assim como hospitalização e morte dos pacientes (77). As vacinas podem evitar a transmissão por dois mecanismos – prevenindo a infecção e tornando os casos menos infecciosos. Porém, conforme descrito na figura, há diversos fatores que contribuem para a eficácia das vacinas, questões como tempo de proteção da vacina, diminuição da transmissão e proteção contra as novas variantes do vírus (77), a Alpha (B.1.1.7) na África do Sul, como a variante Gamma (P.1) no Brasil (92), a Delta e sublinhagens (94) e mais recentemente, a Omicron e sublinhagens

(B.1.1.529). Com a evolução da pandemia, o surgimento das VOCs desafiam a eficácia das vacinas. Estudos posteriores estabeleceram claramente que a eficácia da vacina contra infecções e doenças sintomáticas diminuiu com o tempo (101) de modo que doses adicionais de vacina (reforços) podem ser benéficas, particularmente para indivíduos mais velhos (102). Mesmo com a vacinação generalizada, os níveis de transmissões continuam altos e o sucesso evolutivo das variantes de SARS-CoV-2 até agora, evoluíram com uma maior transmissibilidade em relação às variantes ancestrais. Variantes como a Omicron e suas sublinhagens, ganham muito de sua vantagem de transmissão ao fugir da imunidade, podem se tornar a norma, como é o caso da gripe sazonal (77).

1.14 SARS-CoV-2 no Brasil

Os primeiros casos na América Latina foram relatados em 26 de fevereiro de 2020 em São Paulo - Brasil, de um viajante que retornava da Itália (103), seguido do Equador (104) e da Argentina (105). Ainda na primeira quinzena de março de 2020, foi confirmada a primeira morte por COVID-19 no estado de São Paulo e todos os estados brasileiros relataram pelo menos um caso confirmado da doença. Nesse mesmo período, o Ministério da Saúde do Brasil declarou surto de transmissão comunitária em larga escala do vírus, onde através de análises filogenéticas posteriores, foi possível inferir diversas introduções virais no Brasil, colaborando assim para a disseminação do vírus dentro do Brasil (91,106).

O governo brasileiro introduziu medidas de restrição na tentativa de mitigar a propagação do vírus. As intervenções não farmacêuticas (NPIs) incluíram bloqueios, medidas de distanciamento social, seguido por fechamento de escolas, universidades e comércios não essenciais. Medidas adicionais incluíram a obrigatoriedade do uso de máscaras de proteção individual, cancelamento de eventos e a abertura apenas de serviços considerados essenciais, como mercados e farmácias (107). A adoção de NPIs durante a primeira onda da pandemia serviu para achatá-la, ou seja, estender o período de ocorrência dos casos. Achatá-la era desejável por três razões: primeiro, evitaria que o sistema de saúde ficasse sobrecarregado

porque o número máximo de leitos ocupados a qualquer momento seria menor sob uma curva mais plana; em segundo lugar, retardaria o ímpeto do surto, reduzindo a superação de casos após o pico do surto e terceiro, permitiu tempo para melhorar as estratégias e a capacidade de atendimento clínico e avaliar a terapêutica (77).

Antes do COVID-19, o sistema de saúde brasileiro já sofria devido a surtos anteriores causada por arboviroses como DENV (Dengue Vírus), ZIKV (Zika Vírus), CHIKV (Chikungunya Vírus) e FAV (Febre Amarela Vírus). Escassez de testes de diagnósticos disponíveis, equipamentos de proteção individual (EPI), falta de preparação e respostas descoordenadas, geradas por um sistema de saúde frágil e falho, baixo acesso ao sistema de saúde, além da pobreza, culminou em tumultos e resultou em subnotificação de casos de SARS-CoV-2, elevado número de mortes e colapso no sistema de saúde, algo não só observado no Brasil como em toda América Latina (108,109). Em maio de 2020, a região Latina se tornou o novo epicentro da pandemia, sendo o Brasil, o segundo país no mundo com maior número de casos confirmados e de mortes registradas (110). As recomendações de distanciamento social e de higiene foram certamente dificultosas nas regiões mais precárias e pobres das grandes cidades brasileiras, como São Paulo e Rio de Janeiro, revelando uma subnotificação de casos e de mortes muito maior do que divulgado pelos órgãos de saúde (111).

Mesmo com o aumento de casos aumentando em todos os estados brasileiros, as medidas de restrição foram flexibilizadas de forma progressiva ao longo dos meses para mitigar os impactos negativos na economia. As viagens de curta duração foram importantes para a transmissão de casos nas cidades fora dos grandes centros e entre estados brasileiros (106), contribuindo para o aumento de casos e mortes de norte (112) a sul (113) e culminando na segunda onda a partir de Janeiro de 2021, ocasionada pela VOC Gama. Nesse período, o Brasil volta a ser o epicentro da pandemia devido ao expressivo número de casos de COVID-19 em todo país, mesmo após a introdução da vacinação nos grupos prioritários, alterando o padrão da faixa etária de infectados, sendo jovens e adultos, os mais infectados nesse período.

No final do primeiro semestre de 2021, a linhagem Delta, outra VOC identificada, foi a responsável pela maioria dos infectados. Esse período foi

marcado por níveis mais baixos de transmissão e menor incidência, provavelmente devido à imunidade adquirida por infecção anterior somado a cobertura vacinal (114). Atualmente, a linhagem VOC predominante no Brasil é a Omicron/VOC, que está em circulação desde final de Dezembro de 2021 (115). Verificou-se que a Omicron/VOC tem uma chance cinco vezes maior de reinfeção em comparação com Delta (116) e apresenta melhor evasão imunológica. Essas características permitiram que essa linhagem seja responsável por reinfeções e devido a sua disseminação contínua, o surgimento de centenas de sublinhagens, inclusive, com recombinação (115).

1.15 Epidemiologia de doenças infecciosas

Os objetivos da epidemiologia viral é compreender a transmissão e como controlar determinado patógeno. Vários fatores podem influenciar a transmissão do vírus, incluindo incidência, prevalência, modo de transmissão do patógeno, suscetibilidade, densidade populacional e padrões de mobilidade (117). A taxa de transmissão de um vírus em uma determinada população também é afetada por fatores biológicos derivados do contato com seu hospedeiro, e por quatro parâmetros importantes: A- o período de latência, definido como o tempo entre a infecção até se tornar infeccioso; B- a variabilidade individual da infecciosidade, definida pela transmissibilidade intrínseca e taxa de contato; C- período de incubação, o tempo entre a infecção e início dos sintomas e D- o intervalo seriado, definido como o tempo entre o início dos sintomas em um par de infectante e infectado (118), conforme ilustrado na **Figura 8**.

Figura 6. Parâmetros epidemiológicos na transmissão de SARS-CoV-2

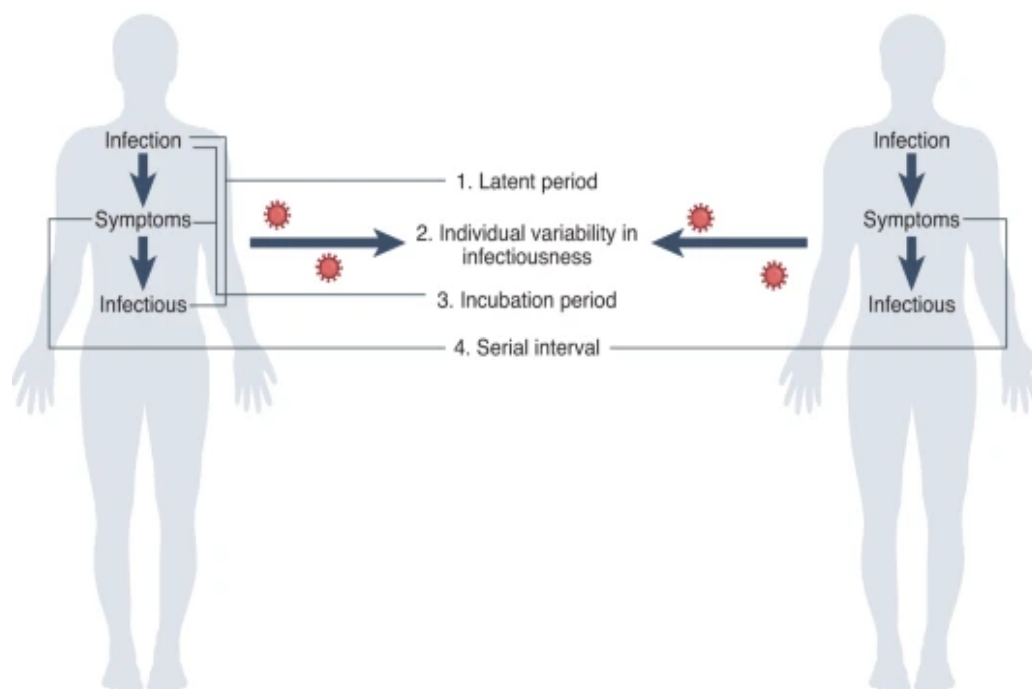


Figura. A- Período latente; B- variabilidade individual na infectiosidade; C- período de incubação e D- intervalo seriado (118).

Um dos parâmetros utilizados para compreensão da extensão da transmissão de um patógeno, termo muito utilizado no início da pandemia ou do início de uma nova variante - é o número de reprodução básico, R_0 , que é definido como a capacidade de contágio de um determinado patógeno sem qualquer tipo de intervenção, representado pelo número médio de casos secundários derivados de um único indivíduo infeccioso em uma população suscetível, onde um $R_0 > 1$ significa uma epidemia crescente, enquanto um $R_0 < 1$ significa a epidemia está diminuindo. Esse dado é usado frequentemente para entender o quão rápido esse patógeno pode se espalhar em uma população; a magnitude que uma epidemia pode atingir, extensão da mitigação de estratégias necessárias para controlar sua disseminação, por exemplo, definir qual grupo será prioritário em campanhas de vacinação para prevenção de novos surtos (56,119). Quando se tem uma epidemia em curso, é necessário compreender o R_e ou R_t , velocidade de disseminação de determinado patógeno ao longo do tempo, influenciada através de medidas e ações que são implementadas (como uso de máscaras, distanciamento social

e aumento da imunidade após infecção ou vacinação, por exemplo) (8). Esses valores de R_0 e R_t são heterogêneos em função dos parâmetros citados anteriormente e têm sido muito utilizados na epidemia atual (60,120). Dados incompletos, subnotificação de casos, falta de padronização na coleta de dados e limitação são exemplos de vieses e que podem inferir análises incorretas. (118).

1.16 Epidemiologia genômica

As sequências genômicas fornecem informações importantes que podem ser utilizadas em conjunto com os métodos tradicionais. Pela primeira vez, o genoma sequenciado pode ajudar a orientar a resposta da saúde pública a uma pandemia quase em tempo real, colaborando na investigação de sua disseminação global, investigação da dinâmica de surtos, propagação espaço-temporal, vias de transmissão e colabora com o monitoramento e identificação de linhagens emergentes. A análise dos genomas do vírus SARS-CoV-2 pode, portanto, complementar, aumentar e apoiar estratégias para reduzir o impacto da COVID-19 (121,122).

A primeira vez que o sequenciamento foi utilizado durante um surto, foi em 2002-2003 durante a epidemia de SARS-CoV, porém, com pouquíssimas sequências disponibilizadas, devido a capacidade limitada de sequenciamento e atraso na divulgação das sequências (121) impactaram na implementação de estratégias para reduzir a disseminação do vírus. Nesse mesmo período, a plataforma online GISAID (Avian Influenza Data) (123) foi criada para facilitar o compartilhamento dessas sequências. Estudos posteriores foram essenciais para determinar a origem da doença (124). Mas foi durante o surto do vírus Ebola (EBOV) ocorrido na Guiné, Libéria (125) e Serra Leoa (126) que os primeiros estudos de epidemiologia genômicos foram realizados realmente em tempo real. As análises filogenéticas e filogeográficas foram utilizadas para identificação da disseminação do vírus em diferentes regiões geográficas e importantes para projetar medidas de controle de doenças mais eficazes (127).

Foi também na epidemia do Ebola que foi utilizado o sequenciador portátil para geração de dados genômicos em tempo real (128). Após essa revolução, o sequenciador portátil foi utilizado em diversos trabalhos

importantes de vigilância genômica como o do Zika vírus (129), assim como, identificar a substituição do genótipo asiático pelo genótipo ECSA (East-Central-South-African) no Chikungunya (130–132), a identificação de um novo genótipo de Dengue 2 associado a uma grande quantidade de casos na região sudeste do Brasil em 2019 (133); nos estudos Febre Amarela no estado de São Paulo (134), onde se confirmou que casos estavam sendo sustentadas por eventos de transbordamento entre primatas não humanos (PNH) e mosquitos vetores silvestres e mais recentemente, na caracterização dos primeiros casos de SARS-CoV-2 no Brasil (103) e descrição de 425 genomas de SARS-CoV-2 provenientes de 22 estados do Brasil (91), na identificação das VOCs Gamma (P1) em Manaus (92) e dos primeiros caso da Beta/VOC identificados (B.1.1.7) no Brasil (86).

1.17 Filogenia, filodinâmica e filogeografia

Muitas aplicações importantes em relação à genômica dos vírus foram baseadas em análises filogenéticas ou filodinâmicas. As análises filogenéticas permitem inferir uma árvore que representa as relações entre os genomas virais estudados e apresenta ramos medidos em unidades de tempo (135) e essas árvores filogenéticas são visualizações úteis de tais relacionamentos. Os padrões de ramificação e o comprimento dos ramos podem ser usados para representar parentesco evolutivo, onde é possível avaliar um grande número de árvores para determinar qual é mais provável de representar a verdadeira relação entre esses organismos (121). Esses ramos indicam os eventos de transmissão de um caso a outro e podem explicar, como o surto se desdobra e utilizar essas informações para a identificação de cadeias de transmissão do vírus(136). Essas informações de inferências filogenéticas baseadas em dados genômicos podem também, estimar uma possível taxa de transmissão do vírus dentro de uma população através da integração de dados genômicos e de metadados (135). Para inferir árvores filogenéticas é necessário selecionar um modelo de substituição adequado para seu banco de sequências, para isso diversos modelos matemáticos foram desenvolvidos para reconstruir os processos de substituição dos nucleotídeos de uma sequência genética a partir

de uma sequência de referência. As ferramentas estatísticas JModelTest (137) e/ou ModelFinder (138) podem ser utilizadas para identificação do modelo mais adequado para o conjunto de dados, colaborando assim, numa inferência filogenética mais adequada. Para estudar a dinâmica temporal de epidemias de vírus, modelos de relógio molecular são normalmente usados para o entendimento das epidemias virais, as datas de coleta das amostras usadas para sequenciamento de genoma tornaram-se uma peça essencial de metadados para estimar árvores filogenéticas calibradas no tempo e para auxiliar no enraizamento filogenético adequado (139,140), sendo utilizado para identificação de novos vírus ou de linhagens novas em uma região específica, como no estudo de Zika Vírus no Brasil(129) e estimar data de transbordamento de eventos, conforme o estudo sobre a infecção entre camelo-humano na infecção por (141) e análise genômica e epidemiológica de Febre Amarela no Brasil(142), por exemplo. Filogenias datadas são importantes para estimar Re , impacto das intervenções e fatores associados à dispersão ao longo do tempo (143–145).

Para discutir sobre análise filodinâmica (144), é necessário diferenciar taxa de mutação e taxa evolutiva (ou taxa de substituição). A taxa de mutação é o número de erros que ocorrem na cópia do genoma durante a replicação. A maioria das mutações é deletéria (44) e virions individuais contendo tais mutações muitas vezes não conseguem se replicar, somente as mutações que aumentam em frequência e se fixam dentro de uma linhagem, de forma aleatória ou devido a ação da seleção natural contribuem para a taxa evolutiva. Então, a taxa evolutiva é descrita como o número de nucleotídeos substituídos por *site*, por ano (muitas vezes abreviado para subs/*site*/ano). A taxa evolutiva muitas vezes pode ser inferida diretamente dos dados da sequência genômica do vírus obtidos de diferentes pacientes em diferentes datas. As linhagens de SARS-CoV-2 acumulam diversidade genética ao longo semanas ou meses em vez de dias, para que os pacientes vizinhos em uma cadeia de transmissão podem ser infectados por vírus com genomas idênticos (121).

O estudo da filogeografia analisa os processos da propagação geográfica através das análises de padrões espaciais das filogenias (146). É possível que o mesmo vírus apresente dinâmicas diferentes dependendo da

escala espacial. Essa dinâmica reflete as taxas relativas de fluxo gênico do vírus, por quanto tempo os vírus têm sido associados a populações humanas e seu modo de transmissão (147). As localizações geográficas são as características mais comumente usadas em filodinâmica e várias abordagens estatísticas foram desenvolvidas para a análise de conjuntos de dados georreferenciados (148). Essas análises são importantes principalmente nos estudos de vírus do tipo RNA, como o SARS-CoV-2, já que a capacidade de adquirir, acumular e transmitir mutações em seus genomas podem ser observadas praticamente em tempo real, nesse momento, num processo de análise microevolutiva (145) colaborando para “ insights” epidemiológicos obtidos a partir da reconstrução da dinâmica espacial e temporal e colaborando para respostas e tomadas de decisões na saúde pública (149) e tem o potencial de de compensar os pontos cegos da vigilância e entender melhor os surtos mal amostrados (135)

2. OBJETIVOS

2.1 OBJETIVO GERAL

Investigar a epidemiologia genômica do vírus SARS-CoV-2 circulante na cidade de São Caetano do Sul, região metropolitana de São Paulo, Brasil no período do estudo

2.2 OBJETIVO ESPECÍFICO

Descrever as linhagens de SARS-COV-2 circulantes no primeiro ano da pandemia na cidade de São Caetano do Sul.

Caracterização e correlação dos sintomas notificados à Plataforma São Caetano às linhagens identificadas

Discutir sobre amostragem oportunista da coorte da Plataforma São Caetano e o impacto em estudos a nível de cidade

2.3 MOTIVAÇÃO

A maioria dos estudos epidemiológicos e de caracterização clínica de COVID-19 foram descritas, principalmente, em estudos de coorte em hospitais, com viés para doenças graves. Estudos clínicos baseados na comunidade e em larga escala são difíceis de estabelecer.

3. MATERIAIS E MÉTODOS

3.1 Características do município de São Caetano do Sul (SCS)

O município de São Caetano do Sul (SCS) está localizado na região metropolitana do estado de São Paulo, pertence à grande região do ABC Paulista (compreende os municípios de Diadema, Mauá, Rio Grande da Serra, Ribeirão Pires, Santo André, São Bernardo do Campo, São Caetano do Sul). Possui cerca de 162.000 habitantes em uma área de 15.331km². A apenas 12 km da cidade de São Paulo, seu território faz divisa com a capital paulista (norte e oeste), São Bernardo do Campo (sul e oeste) e Santo André (sul e leste). É intensamente conurbada com os municípios de São Paulo, Santo André e São Bernardo do Campo, fazendo com que os limites físicos dessas cidades sejam perdidos (150,151). O município apresenta IDH alto (0,862) e, baixo índice de analfabetismo (1,5%) em relação ao Estado de São Paulo (4,2%) e demais municípios da Grande São Paulo. A população é 100% urbana e quase 100% dos domicílios contam com infraestrutura básica, como água encanada e coleta de esgoto. A Rede Municipal de Saúde possui uma infraestrutura bem estruturada, contendo 12 Unidades Básicas de Saúde (UBS) com Equipes de Saúde da Família; 5 Centros de Idosos; 15 Unidades Especializadas; 3 hospitais públicos e 1 UPA (Unidade de Pronto Atendimento).

3.2 Programa de atenção Primária Corona São Caetano

Implementado em 6 de abril de 2020, o programa Corona São Caetano (<https://coronasaocaetano.org>) é uma iniciativa de atenção primária que visa organizar a resposta de saúde pública à pandemia de COVID-19. A plataforma foi desenvolvida em parceria com a Faculdade de Medicina da Universidade Municipal de São Caetano do Sul (USCS), Universidade de São Paulo (USP) e colaboração do Instituto de Medicina Tropical da Universidade de São Paulo (IMTSP). O programa orientou os munícipes a se cadastrarem por meio de plataforma digitais (online ou contato telefônico) para receberem uma consulta de telessaúde para avaliação de risco. Mulheres grávidas ou pacientes com sintomas alarmantes, como falta de ar, febre persistente, confusão ou letargia,

foram orientados a procurar atendimento médico no hospital. Os pacientes sem risco foram orientados a permanecer em casa e realizar auto-coleta de amostras para diagnóstico, recebendo atendimento por telessaúde por até 14 dias após o início dos sintomas. O programa visava evitar deslocações desnecessárias às Unidades de Saúde, detectar precocemente doentes infectados e prevenir a propagação do vírus pela cidade (152).

3.3 Seleção representativa de amostras para sequenciamento

A plataforma São Caetano rotineiramente coleta endereços de pacientes e os categoriza em bairros (n=15). Para obter uma representação espacial abrangente dos casos positivos em toda a cidade, selecionamos aleatoriamente duas amostras positivas por semana epidemiológica e por bairro. Essa abordagem nos permitiu incluir 1.063 casos positivos, abrangendo o período de 6 de Abril de 2020 a 30 de Abril de 2021, e garantiu a captura da diversidade de casos de diferentes locais da cidade.

3.4 Aspectos éticos

O projeto intitulado “A Diversidade Genética do Coronavírus no Brasil”, da Pesquisadora Responsável: Ester Cerdeira Sabino, tem a aprovação da Comissão Nacional de Ética em Pesquisa (CEP/CONEP) sob protocolo de número: CAAE 30127020.0.0000.0068 em 23/03/2020 e Comitê de Ética para Análise de Projetos de Pesquisa da USCS (protocolo nº 13.915, de 3 de junho de 2020). Os comitês dispensaram a necessidade de consentimento informado e permitiram o desenvolvimento de um conjunto de dados analíticos não identificados para análise.

3.5 Diagnóstico de SARS-CoV

Todas as amostras coletadas foram encaminhadas ao Laboratório de Virologia ou Laboratório de Parasitologia Médica do Instituto de Medicina Tropical da Universidade de São Paulo e mantidas sob refrigeração (- 20 ° C ou - 80 ° C) até o momento da realização dos testes. A extração do material

genético foi realizado através dos kits QIAamp RNA Blood Mini Kit ou Extracta Kit Fast – DNA e RNA Viral (Loccus) e o diagnóstico foi realizado utilizando kits disponíveis: Kit ALTONA RealStar® SARS-CoV-2 RT-PCR 1.0 (Hamburgo, Alemanha), kit Mico BioMed RT-qPCR (Seongnam, Coréia do Sul) de acordo com as instruções do fabricante.

3.6 Amplificação viral

Para a amplificação do genoma das amostras positivas, foi utilizado o protocolo de PCR Multiplex estabelecido em nosso laboratório. Primeiramente, o cDNA foi sintetizado a partir de amostras de RNA positivas, usando o kit de síntese de cDNA ProtoScript II First-Strand e *random primers* ou LunaScript RT SuperMix Kit (New England Biolabs, Reino Unido). Posteriormente, o cDNA foi submetido à amplificação do genoma completo usando os esquemas de iniciadores ARTIC-V3, V3.1 e V.4 ((153)), Q5 High-Fidelity DNA polimerase ou Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Reino Unido) (154,155).

3.7 Sequenciamento de genoma

As amostras amplificadas foram então transferidas para um tubo do tipo *Lobind* de 1,5ml para purificação, utilizando 1x Ampure XP beads²⁸ (Beckman Coulter, United Kingdom) e solução de álcool 70%, por duas vezes. A eluição foi feita com Água Livre de Nuclease. Após a etapa de purificação, os *amplicons* foram quantificados com fluoróforo Qubit dsDNA High Sensitivity Assay Kit e dosadas no aparelho Qubit 3.0 instrument (Life Technologies, USA (ThermoFisher) conforme recomendação do fabricante. Os *amplicons* de cada amostra foram então normalizadas, deixados em *pool* e identificadas com código de barras (*barcodes*). Para o preparo da biblioteca utilizamos um input de 240ng e dois esquemas diferentes de amostras: I) 23 amostras e 1 controle negativo ou II) 95 amostras e 1 controle negativo, com os seguintes kits: EXP-NBD104 (código de barras 1-12) e EXP-NBD114 (código de barras 13-24) ou EXP-NBD 196 (código de barras 1-96) Native Barcoding Kits e Ligation

²⁸ pérolas magnéticas

Sequencing kit (ONT, SQK-LSK-109) de acordo com o protocolo citado. Finalmente, 30ng da biblioteca final contendo 23 amostras e 1 controle negativo ou 95 amostras e 1 controle negativo, foram carregadas em uma célula de fluxo R9.4.1 no dispositivo MinION/GridION usando os programas MinKNOW (v.22.3.6) e RAMPART para o monitoramento da corrida de sequenciamento em tempo real (<https://artic.network/rampart>). A corrida foi iniciada e teve uma média de duração de 12 horas.

3.8 Montagem do genoma

Os arquivos de saída Fast5 foram submetidos ao programa Guppy versão 6.0.7 (Oxford Nanopore Technologies, Reino Unido) para realização da chamada de bases e separação dos códigos de barras (*basecall* e *demultiplex*). Os arquivos FastQs foram mapeados contra o genoma de referência do SARS-CoV-2 isolado em Wuhan-Hu 1 (GenBank MN908947) para obtenção das sequências consensos com o programas Minimap 2 versão 2.28.0 (156) e convertidas em formato BAM utilizando SAMTools (157). O programa ARTIC guppy-plex foi utilizado para remoção das leituras de baixa qualidade e as variantes foram detectadas através do programa Nanopolish. Todos os detalhes do processo de sequenciamento e análise de bioinformática estão descritos no endereço da web (153). A identificação das linhagens/clados de SARS-CoV-2 foram realizadas através do programa Pangolin (158) e Nextclade (81). As sequências deste estudo foram submetidas ao teste de recombinação usando todos os métodos disponíveis no RDP4 (159) e nenhuma evidência de recombinação foi encontrada.

3.9 Controle de qualidade de sequências de consenso do genoma

As corridas com controle negativo apresentando qualquer contaminação foram descartadas. As regiões do genoma com profundidade <20 vezes não foram incluídas nas sequências de consenso finais e essas posições são representadas com N caracteres. As sequências consensos com cobertura <75% de cobertura do genoma foram removidas das análises.

3.10 Conjuntos de dados genômicos e alinhamento de sequência

Para contextualizar as sequências geradas neste estudo, foi realizado análises filogenéticas, para esse fim, foi construído conjunto de dados de referência global abrangente que incluiu uma sequência Internacional por país por semana epidemiológica, entre Dezembro de 2019 e Abril de 2021 (n = 3460), compreendendo todo o período do estudo. As sequências do genoma de referência foram recuperadas do banco de dados GISAID EpiCoV (123). Além disso, foi adicionado sequências brasileiras de todos os 27 estados federais brasileiros, em quantidades proporcionais ao número semanal de casos relatados (n = 6637; dados de incidência recuperados de <https://brasil.io/dataset/covid19/files/>), conforme descrito em (91,160). O conjunto de dados final, incluindo sequências aqui geradas, compreendeu 10970 sequências. As sequências foram alinhadas com os programas Minimap 2 v2.28.0 (156) e Gofasta v1.2.0 (161). O programa AliView v1.26 (162) foi usado para edição das sequências, retirando as regiões genômicas não traduzidas e visualização do alinhamento final.

3.11 Análise filogenética e sinal temporal

Uma árvore de máxima verossimilhança (ML) foi inferida usando o programa IQTREE v.2(163) utilizando o modelo de substituição de nucleotídeos GTR + F + I + G4(164). O teste de razão de verossimilhança aproximada semelhante a Shimoidara-Hasegawa (SH-aLRT)(165) foi usado para quantificar a incerteza filogenética ao longo dos galhos das árvores, e o TempEST v1.5.1(166) foi usado para avaliar o tempo sinal do conjunto de dados por meio de uma regressão raiz-a-ponta que otimiza a relação entre datas de coleta de amostras e identificar amostras com sinal temporal inconsistente (*outliers* na regressão raiz-a-ponta). Conforme descrito anteriormente(160), as sequências apresentando resíduos de mais ou menos do que 1,5x o intervalo interquartil da distribuição de resíduos foram tratadas arbitrariamente como *outliers* e, portanto, removidas do conjunto de dados.

3.12 Reconstruções filogeográficas bayesianas

Para obter uma grande árvore filogenética em escala de tempo, usamos um modelo BEAST desenvolvido recentemente (BEAST Thorne, BEAST v.1.10.5, pré-lançamento) (167) que reduz a carga computacional durante a análise em relação ao procedimento padrão que usa um modelo de Poisson para dimensionar as distâncias genéticas em uma árvore fornecida ao longo do tempo (168). Esta análise foi realizada usando uma taxa evolutiva fixa de $7,5 \times 10^{-4}$, consistente com estimativas anteriores (169). O *skygrid* demográfico não paramétrico flexível anterior foi empregado, com um ponto de corte definido no final de dezembro de 2019, consistente com a estimativa do eixo x da regressão raiz-a-ponta. O número de grades foi definido para corresponder ao número de semanas compreendidas no intervalo de tempo da árvore ($n = 75$). Esta análise foi executada em dez execuções independentes de Markov Chain Monte Carlo (MCMC), com 500 milhões de gerações, definidas para amostrar a cada 25.000 estados. O programa Tracer v.1.7.1 (170) foi usado para verificar a mistura e convergência (tamanho efetivo da amostra >200 para todos os parâmetros) e Logcombiner v.1.10.4 (167) foi usado para mesclar todos os *logs* com 10% de queima. O programa também foi usado para amostrar 1.000 árvores datadas da distribuição posterior.

As análises filogeográficas foram então realizadas neste conjunto de árvores datadas empíricas (91). Usamos modelos filogeográficos assimétricos discretos (149) e um método de contagem robusto (171,172) para mapear todas as transições entre locais ao longo da árvore. O esquema de discretização considera um modelo com quatro categorias: Internacional ($n = 3.460$), Brasil (exceto estado de São Paulo; $n = 5.197$), estado de São Paulo (exceto município de São Caetano, $n = 1.440$) e município de São Caetano ($n = 873$). Os programas Tracer v.1.7.1(167,170) e Logcombiner v.1.7.1 foram utilizados e uma árvore de máxima credibilidade de clado foi inferida com TreeAnnotator v.1.10.4 (167).

3.14 Dados epidemiológicos e clínicos

Para realizar análises epidemiológicas e clínicas, extraímos dados da plataforma Corona São Caetano para casos confirmados de SARS CoV-2. Foram coletados dados clínicos não identificáveis, incluindo informações demográficas (idade, sexo, nível de escolaridade, ocupação e bairro) e informações clínicas, como data de início dos sintomas e sintomas relatados, coletados através de um questionário padronizado (**arquivo complementar 2S**). Os dados de Síndrome Respiratória Aguda Grave (SRAG) e internações foram extraídos da plataforma SIVEP-Gripe (Sistema de Informação de Vigilância Epidemiológica da Influenza), que pertence ao banco de dados do Ministério da Saúde. Todos os dados são de pessoas físicas que têm como residência a cidade de São Caetano do Sul. Também foram incluídos dados do Boletim Municipal.

3.15 Inferência da prevalência da linhagem ao longo do tempo

A prevalência diária de cada linhagem foi estimada aplicando uma janela móvel de tamanho variável ao número diário de amostras sequenciadas identificadas como aquela linhagem. A cada instante, apenas as amostras contidas na janela móvel são consideradas no cálculo da prevalência. O algoritmo utiliza um tamanho de janela maior em períodos onde poucas amostras são sequenciadas, evitando ruído excessivo e diminuindo o tamanho da janela em períodos com grande número de sequências, aumentando a precisão do tempo. Optamos por usar uma janela variável no tempo porque algumas linhagens não atingem uma prevalência próxima a 100%, portanto, regressões logísticas multinomiais como as usadas em (173) não podem ser usadas para inferir a prevalência em tempo contínuo.

Seja $S[n, l]$ o número de amostras sequenciadas associadas à linhagem $l \in L$ no dia n , onde $L = \{\text{Gamma}, \text{Zeta}, \text{B.1.1.28}, \text{B.1.1.33}, \text{Outros}\}$. Todas as linhagens que não são classificadas como Gamma, Zeta, B.1.1.28 ou B.1.1.33 são designadas em um único grupo 'Outros'. Observe que $S[n, l]$ é um sinal esparsos para um l fixo, sendo diferente de zero apenas nos dias em que uma amostra sequenciada de PCR+ é identificada como pertencente à linhagem l .

Defina para cada instante $n \geq 1$ o raio da janela $L[n]$ como o menor $L' \geq L_{min}$ que

$$\sum_{m=n-L'}^{n+L'} \sum_{l \in L} S[m, l] \geq N_{min},$$

onde L_{min} and N_{min} são parâmetros predefinidos que representam, respectivamente, o menor raio da janela e o número de amostras permitido. Essas quantidades são arbitrárias, e neste trabalho escolhemos $L_{min} = 7$ dias e $N_{min} = \lceil (2L_{min} + 1) \frac{N_{total}}{\Delta t} \rceil = 34$, onde $N_{total} = 879$ que é o total de sequências produzidas no período e $\Delta t = 391$ dias é a duração do estudo. Portanto, N_{min} é o número médio de amostras contidas em uma janela de raio $L_{min} = 7$ dias. Assim, este método define uma janela cujo raio variável no tempo não pode ser menor que 7 dias e aumenta até que um mínimo de 34 amostras sejam incluídas na janela.

Uma vez que $L[n]$ é computado, a prevalência de cada linhagem l no instante n , é denotado por $\rho[n, l]$, é obtido calculando a prevalência bruta das amostras no intervalo $[n - L[n], n + L[n]]$:

$$\rho[n, l] = \frac{\sum_{m=n-L[n]}^{n+L[n]} S[m, l]}{\sum_{l \in L} \sum_{m=n-L[n]}^{n+L[n]} S[m, l]}$$

3.16 Estimativa do número de reprodução efetiva

Para inferir estimativas do número efetivo de reprodução (R_t) específico da linhagem utilizamos o algoritmo EpiSmooth (174) com passo $\eta=0.1$. Usamos uma distribuição log-normal com log-média 1,09 e log-desvio-padrão 0,72 como distribuição de intervalo de geração com base em intervalos seriais brasileiros medidos na fase inicial da epidemia no Brasil. Utilizamos o número diário de casos notificados na Plataforma SC e a prevalência diária estimada $\rho[n, l]$ (175).

Para estimar o R_t da linhagem-específica e o intervalo de confiança, geramos $N_{samples} = 1,000$ amostras da incidência específica da linhagem diária $I[n, l]$ usando uma distribuição de tamanho binomial $I[n] = \sum_{l \in L} I[n, l]$ (número de indivíduos PCR+ que tiveram seus primeiros sintomas no dia n) e probabilidade $\rho[n, l]$. Para cada amostra, $I[n, l]$, estimamos R_t usando EpiSmooth (174). Em seguida, obtemos a mediana e os intervalos de 95% de credibilidade extraíndo os quantis do conjunto de 1.000 amostras R_t para cada n e l . Também foi estimado o R_t sem desagregação por linhagem simplesmente executando o EpiSmooth usando o número diário de casos.

3.17 Estimativa de probabilidades de sintomas

Consideramos que um paciente com resultado de PCR+ sequenciado infectado por uma determinada linhagem teve um determinado sintoma se o sintoma foi relatado em alguma das consultas médicas. Os intervalos de credibilidade foram calculados usando uma abordagem bayesiana assumindo uma distribuição a priori uniforme para a probabilidade de ter o sintoma em questão p_s no intervalo $[0, 1]$. A probabilidade é $N^+ | p_s \sim Binomial(N, p_s)$, onde N é o número total de pacientes PCR+ que foram sequenciados e infectados por essa linhagem e N^+ é o subconjunto deste grupo que teve o sintoma dado. Portanto, a distribuição a posteriori é $p_s | N^+ \sim Beta(1 + N^+, 1 + N - N^+)$. Os quantis da distribuição Beta são desenhados para estimar a mediana e os intervalos de 95% de credibilidade mostrados na **Figura 5A**.

Para validar nossos resultados, as probabilidades de sintomas também foram estimadas considerando todos os pacientes PCR+, não apenas pacientes com amostras sequenciadas. Para isso, imputamos a linhagem que infectou cada paciente com base na prevalência de linhagem inferida $\rho[n, l]$ no dia do início dos sintomas. O procedimento usado para obter a distribuição posterior para p_s é o mesmo que acima, exceto que N^+ e N são estimados por

$$N = \sum_{i=1}^{N_{PCR^+}} \rho[n_i, l] \text{ e } N^+ = \sum_{i=1}^{N_{PCR^+}} \rho[n_s[i], l] V[i], \text{ onde } N_{PCR^+} \text{ é o número de pacientes}$$

PCR+, n_i é a data de início dos sintomas do paciente i e $V[i] = 1$ se paciente eu tive o sintoma dado, caso contrário $V[i] = 0$.

3.18 Taxas de Fatalidade e Hospitalização

As Taxa de Letalidade (CFR), Taxa de Hospitalização (CHR) e Taxa de Fatalidade Hospitalar (HFR) foram estimadas usando casos notificados na plataforma SC até 31 de março de 2021 e internações por Síndrome Respiratória Aguda Grave (SRAG) relatadas no SIVEP-Gripe. Este é um conjunto de dados aberto que contém informações a nível individual, como cidade de hospitalização, cidade de residência, data de início dos sintomas e desfecho (ou seja, recuperação ou morte). Este conjunto de dados está disponível em <https://opendatasus.saude.gov.br/>. Selecionamos pacientes com SRAG residentes em São Caetano do Sul e internados no estado de São Paulo. Calculamos o número diário de casos e óbitos usando a data de início dos sintomas, disponível em ambos os conjuntos de dados.

Como no procedimento usado para estimar probabilidades de sintomas, usamos uma abordagem Bayesiana para inferir as taxas de fatalidade e hospitalização assumindo uma distribuição uniforme no intervalo $[0,1]$ para a probabilidade de desfecho grave ou fatal p_S . Portanto, a distribuição a posteriori da taxa de letalidade ou internação em determinado período é

$$p_S | N_S \sim \text{Beta}(1 + N_S, 1 + N_R),$$

enquanto, N_S é o número de resultados graves ou fatais observados e N_R é o número observado de recuperações. Como tal, N_S é o número de internações para cálculo do Taxa de Letalidade e o número de óbitos para estimativas de Taxa de Fatalidade Hospitalar e Taxa de Hospitalização, enquanto N_R é o número de recuperações de SRAG para inferência de Taxa de Fatalidade Hospitalar e a diferença entre o número de casos de PCR+ e o número de mortes para cálculos de Taxa de Fatalidade Hospitalar e Taxa.

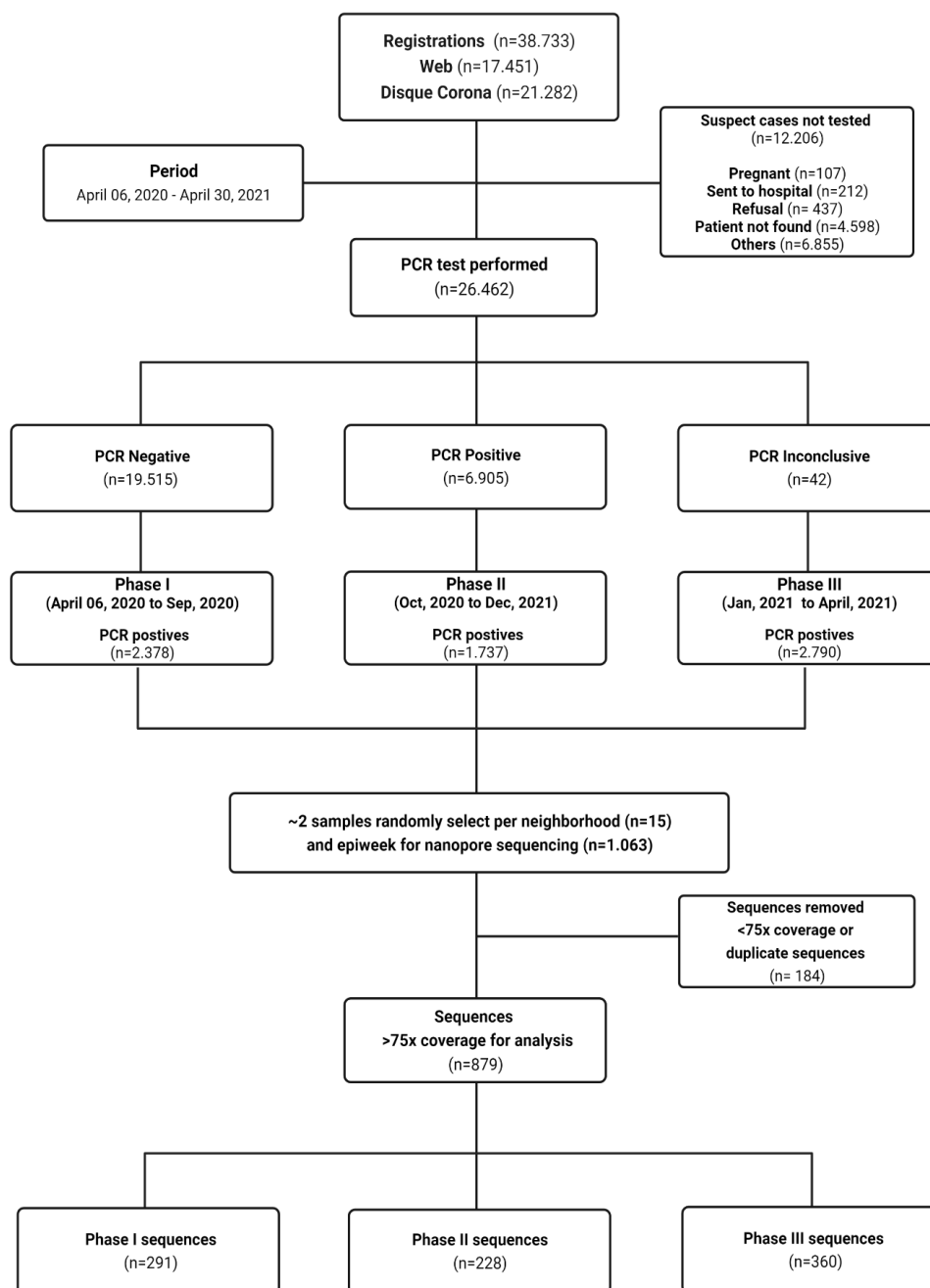
4. RESULTADOS

4.1 Contexto epidemiológico

A epidemia de SCS pode ser dividida em três fases que compreendem: Fase 1: entre Abril e Setembro de 2020, onde ainda há circulação das linhagens ancestrais e não há detecção de VOC/VOI; Fase 2 de Outubro a Dezembro de 2020, quando a linhagem Zeta (P.2) começa a circular e Fase 3 entre Janeiro e Abril de 2021, quando a linhagem Gama (P.1) circulou amplamente no Brasil.

Durante o período do estudo, (6 de abril de 2020 - 30 de abril de 2021), o Boletim Municipal de São Caetano do Sul registrou 10880 casos de COVID-19. Nesse mesmo período, 38733 pacientes com suspeita de COVID-19 foram cadastrados na plataforma Corona São Caetano, dos quais 26584 foram testados por RT-PCR e um total de 6905 pacientes (26,09%) foram positivos, o que corresponde a 63% do total de casos notificados do município. Selecionamos 1063 (15,4%) amostras para sequenciamento e conseguimos obter sequências com mais de 75% de cobertura em 879 amostras (**Figura 7**).

Figura 7. O fluxograma contém informações sobre o desenho do estudo.



Legenda. Durante o período do estudo, 38733 pacientes foram cadastrados na Plataforma SC para atendimento. Dos 26.462 indivíduos que se auto coletaram para o diagnóstico de SARS-CoV-2, um total de 6.905 indivíduos foram positivos, 19.515 foram negativos e 42 foram inconclusivos. A partir dessas amostras positivas, selecionamos aleatoriamente ~ 2 amostras positivas por bairro por semana epidemiológica para sequenciamento do genoma usando o protocolo ARTIC e a plataforma de sequenciamento Oxford Nanopore. Para

análises epidemiológicas e atribuição de linhagem, consideramos genomas com cobertura >75x (n=879).

Na **Tabela 1** descrevemos as características sociodemográficas de todos os casos positivos e casos sequenciados. De forma resumida, observamos a predominância do sexo feminino e de adultos na faixa etária entre 20 e 59 anos. Ambos os grupos foram muito semelhantes em relação à proporção de gênero, idade, escolaridade e exposição.

Tabela 1. Características demográficas e clínicas.

Tabela 1. Características demográficas e clínicas dos pacientes do programa Corona São Caetano					
Características	6905 casos positivos		879 sequências de COVID-19		p-value **
	Valores	Freq (%)	Valores	Freq (%)	
Sexo					1.335
Masculino	2922	42.31%	393	44.71%	
Feminino	3983	57.68%	486	55.29%	
Faixa Etária (%)	12 ≤ 44 ≤ 105		12 ≤ 43 ≤ 92		4.437
12–19	363	5.29%	41	4.66%	
20–39	2618	37.94%	329	37.43%	
40–59	2590	37.54%	349	39.70%	
60+	1334	19.22%	160	18.20%	
Educação					6.166
Nunca foi à escola	63	0.91%	10	1.14%	
Até primário	1284	18.60%	153	17.41%	
Ensino Médio	3065	44.39%	401	45.62%	
Universidade	2467	35.73%	308	35.04%	
NA	26		7		
Ocupação Essencial *					7.329
Trabalhador de áreas essenciais*	1551	22.46%	204	23.21%	
Cuidador	98	1.42%	13	1.48%	
Profissional da Saúde	325	4.71%	46	5.23%	

continuação

Não trabalhador
de áreas
essenciais

	4904	71.02%	609	69.28%
NA	27	0.39%	7	0,80 %

NA= sem
informação

* Trabalhadores de segurança, serviços de emergência, supermercados, transportes públicos e farmácia

** Pearson's Chi-squared test; Fisher's exact test

Das 879 sequências obtidas com cobertura genômica >75%, as linhagens mais comuns detectadas foram B.1.1.28, B1.1.33, Zeta/VOI a e Gama/VOC em 26,2%, 8,7%, 11,7% e 44,1%, respectivamente. Outras variantes foram detectadas nos restantes 9,3% dos casos (B.1.1.332, B.1.1, B.1.1.174, B.1.1.348, B.1.1.135, B.1.1.371, B.1.1. 274, B.1.1.397, B.1.1.332, B. 1.1.10 e Alpha/VOC) (**Tabela 2**).

Tabela 2. Principais grupos de linhagens de SARS-CoV identificadas em São Caetano do Sul, agrupados por fase (I, II, III).

Mês/Ano	SARS-CoV-2 LINHAGENS								Total	FASE
	B.1.1.2 8	B.1.1.3 3	B.1.1.7	N.9	Outros *	P.1 **	P.2	P.7		
abril/2020	59	19			1				79	Fase 1
maio/2020	59	10			4		1		74	
jun/2020	29	16			1				46	
jul/2020	37	6			2				45	
agosto/2020	22	7			3				32	
Setembro/2020	13				1				14	
Outubro/2020	43	4		1	4		2		54	Fase 2
Novembro/2020	34	5		4	4		16	2	65	
Dezembro/2020	57	7		5	13		26	3	111	
Janeiro/2021	27	1		6	8	16	43	3	104	Fase 3
Fevereiro/2021	5	1		2	3	69	6	1	87	
Março/2021	2	1	5			124	9	1	142	

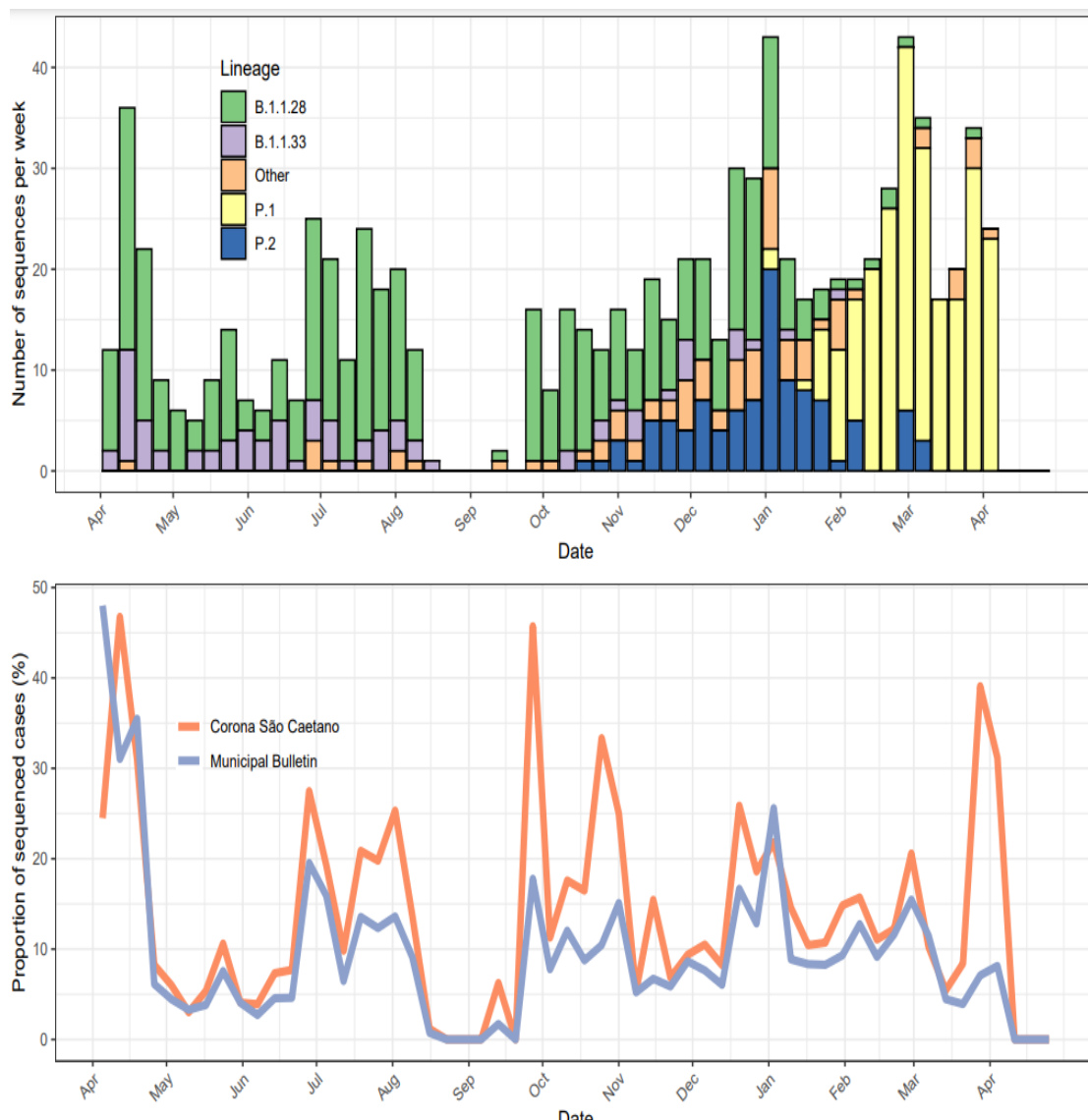
Abril/2021	1	2	1	22	26				
	388	77	7	18	45	231	103	10	879

* Outros: B.1.1.332, B.1.1, B.1.1.174, B.1.1.348, B.1.1.135, B.1.1.371, B.1.1.274, B.1.1.397, B.1.1.332
B.1.1.10

** P.1 grupo : P.1, P.1.14, P.1.10

Descrevemos a evolução das variantes ao longo do tempo e a proporção de casos sequenciados em relação à Plataforma SC e ao Boletim Municipal. As linhagens ancestrais B.1.1.28 e B1.1.33 co-circularam na primeira onda, com casos diminuindo em agosto de 2020, mas se recuperando lentamente quando Zeta VOI começou a se espalhar. O primeiro caso de infecção por Gama (P.1) foi detectado no início de Janeiro de 2021 e, em meados de Fevereiro de 2021, essa linhagem tornou-se a variante predominante e responsável pela segunda onda.

Gráfico 1. Evolução das variantes em São Caetano do Sul, São Paulo.

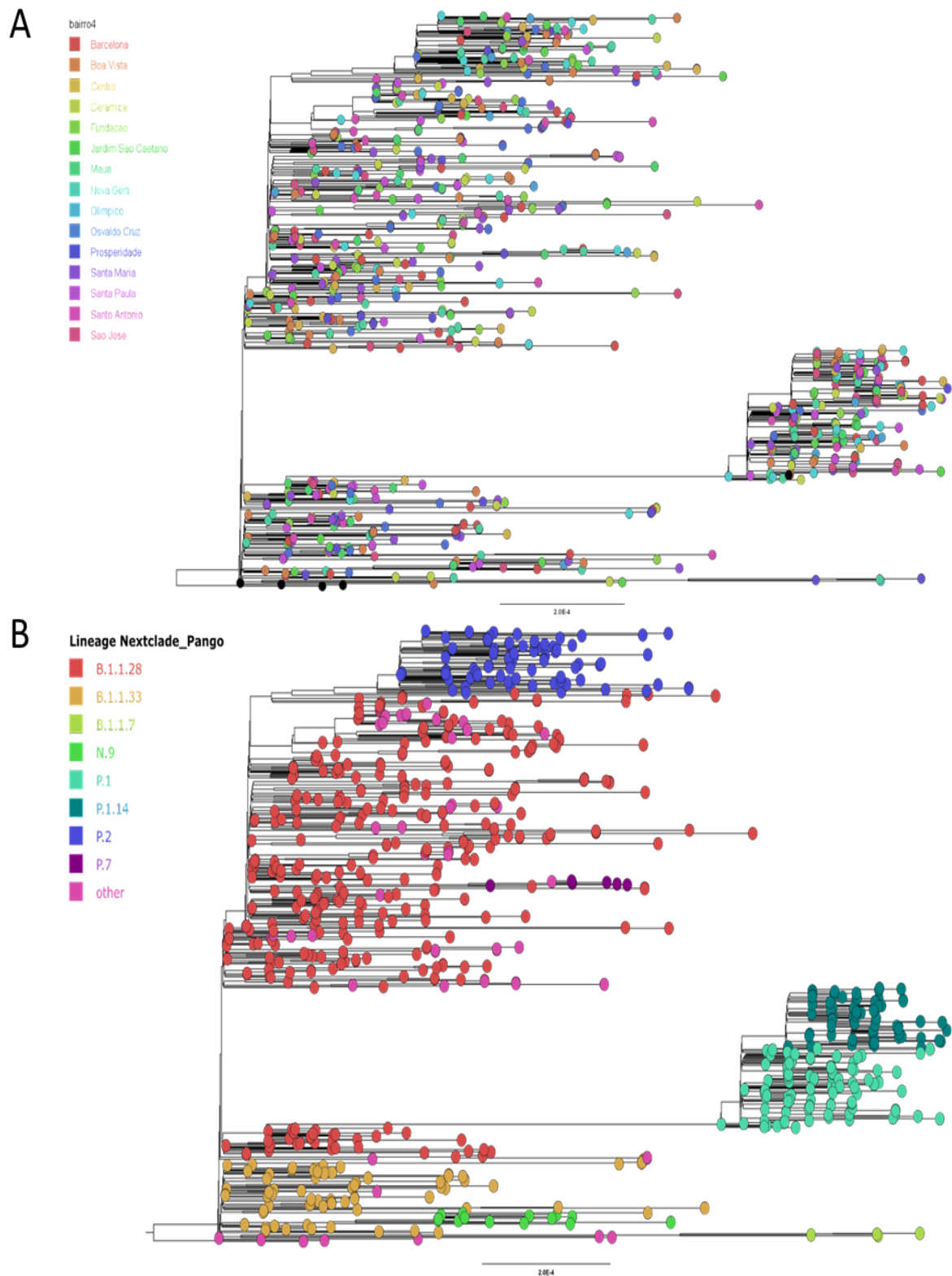


Legenda 10. (A) Número de amostras PCR+ sequenciadas por semana epidemiológica identificadas como as principais linhagens ancestrais (B.1.1.28 e B.1.1.33), VOI (Zeta) e VOC (Gamma). As amostras não identificadas como uma dessas linhagens foram colocadas no grupo 'Outras'. **(B)** Proporção de casos PCR+ sequenciados entre todos os casos notificados pelo boletim municipal (em azul) ou Plataforma SC (laranja).

4.2 Análise filogenética

Foi realizada uma árvore somente com as sequências deste estudo (n=879). Nessas árvores de ML podemos inferir que a disseminação dentro do município ocorreu de forma heterogênea em todas as fases (Figura 8A). Na Figura 8B, classificamos a árvore de acordo com a linhagem atribuída a cada sequência.

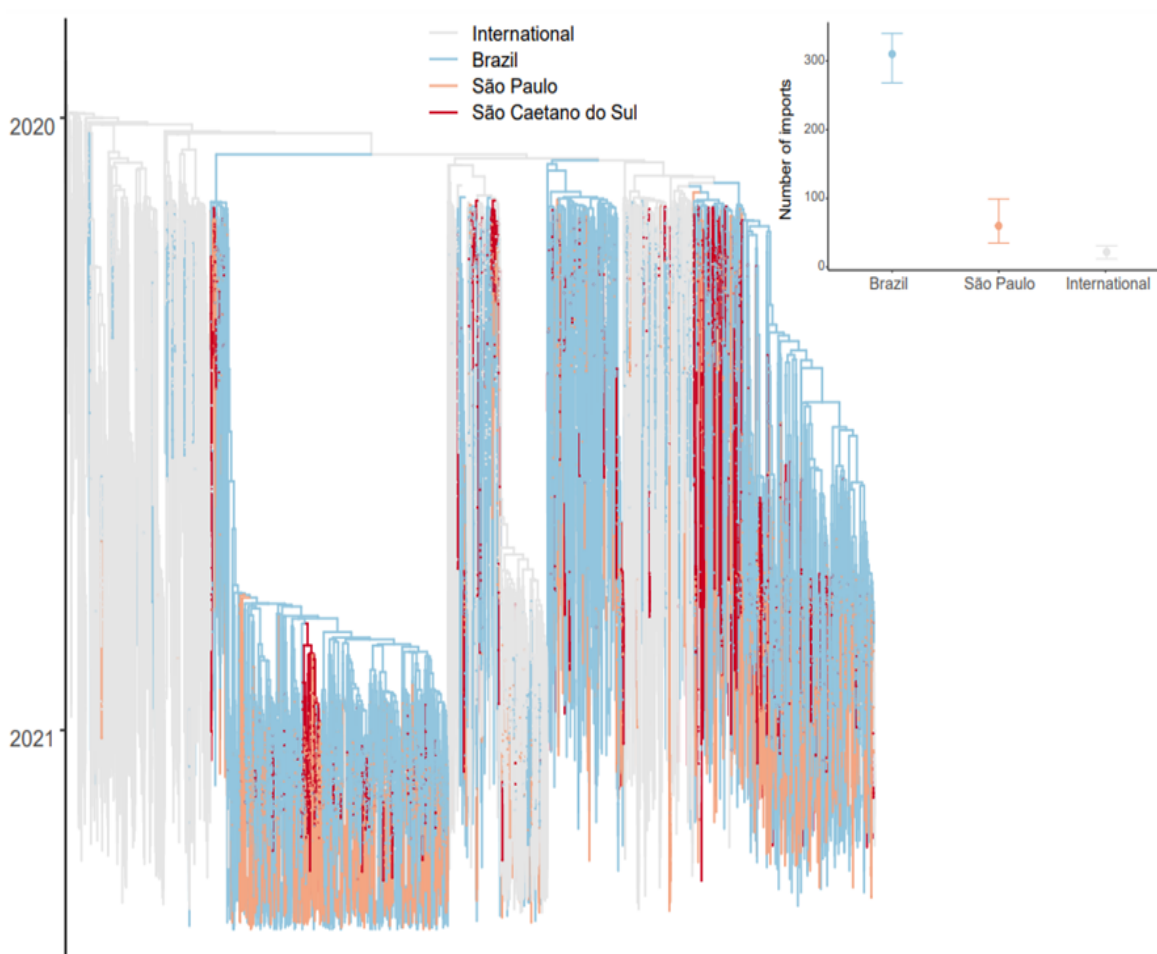
Figura 8. Análise filogenética das amostras de SARS-CoV-2 da cidade de São Caetano do Sul identificadas pelo bairro (A) e linhagem (B).



4.3 Inferências filogeográficas de relógio molecular

Para determinar os padrões de propagação espacial e temporal que levam ao estabelecimento de linhagens virais em São Caetano do Sul, foi realizada uma análise filogeográfica bayesiana com um grande conjunto de sequências, utilizando uma parametrização de menor complexidade. As sequências de São Caetano do Sul agruparam-se com múltiplos clados espalhados na árvore, denotando múltiplas introduções independentes de diversas linhagens ($n=392$), principalmente, de outros estados brasileiros. Introduções internacionais também foram detectadas ($n=26$). Da mesma forma, as exportações de São Caetano do Sul foram detectadas para outros estados brasileiros (mediana: 74, 95% HPD: 42 - 119) e outros municípios do estado de São Paulo (mediana: 154, 95% HPD: 94 - 206).

Figura 9. Inferência filogeográfica

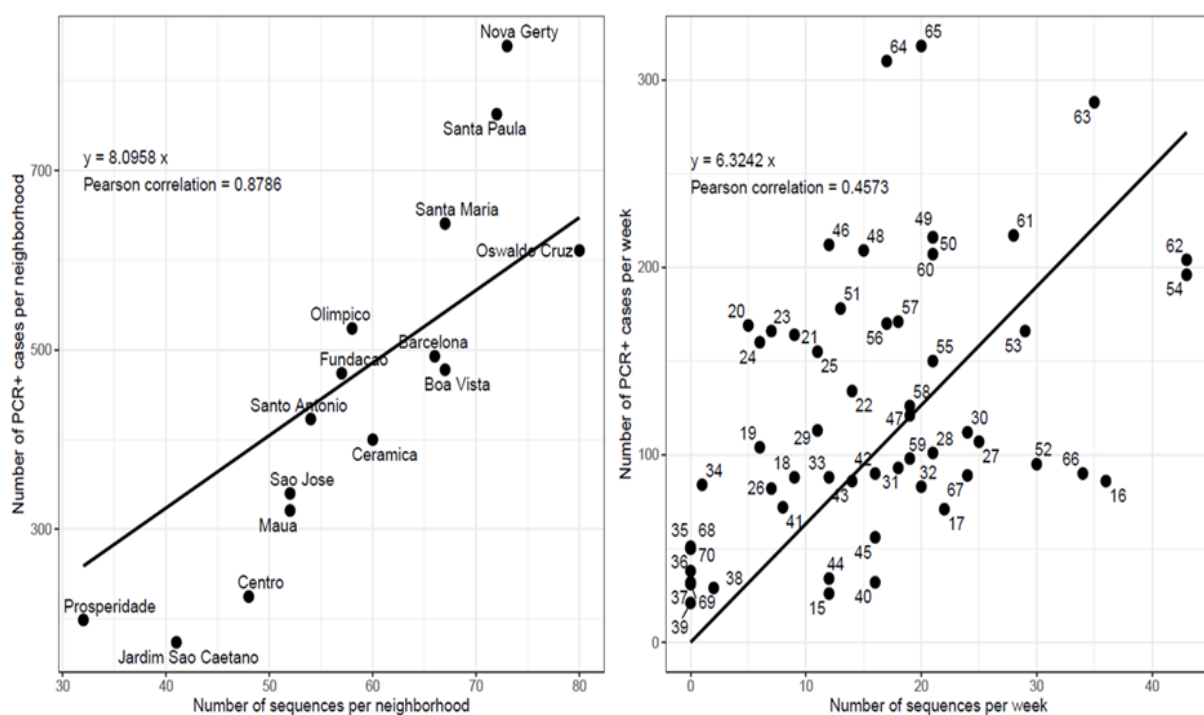


Legenda. O modelo de quatro estados suporta a ocorrência de múltiplas introduções, 60 (95% HPD: 35 - 99) importações de outros municípios do estado de São Paulo; 310

(95% HPD: 268 - 340) de outros estados brasileiros e 22 (95% HPD: 12 - 31) de outros países. Apresentações internacionais para São Caetano do Sul também foram detectadas (mediana: 26, 95% HPD: 11 - 39). Da mesma forma, as exportações de São Caetano do Sul foram detectadas para outros estados brasileiros (mediana: 74, 95% HPD: 42 - 119) e outros municípios do estado de São Paulo (mediana: 154, 95% HPD: 94 - 206).

A estratégia de amostragem para sequenciamento foi essencial na obtenção de um conjunto de seqüências representativas tanto epidemiológico, espacial e temporal. Uma análise de correlação linear usando o Teste de Pearson (**Gráficos 2A e 2B**) foi realizada para comparar os casos de PCR+ por bairro e semana epidemiológica, considerando seqüências com cobertura >75x. Obtivemos bons resultados de correlação tanto espacialmente (Pearson = 0,87) quanto temporalmente (Pearson = 0,45).

Gráfico 2. Correlação do banco de dados.



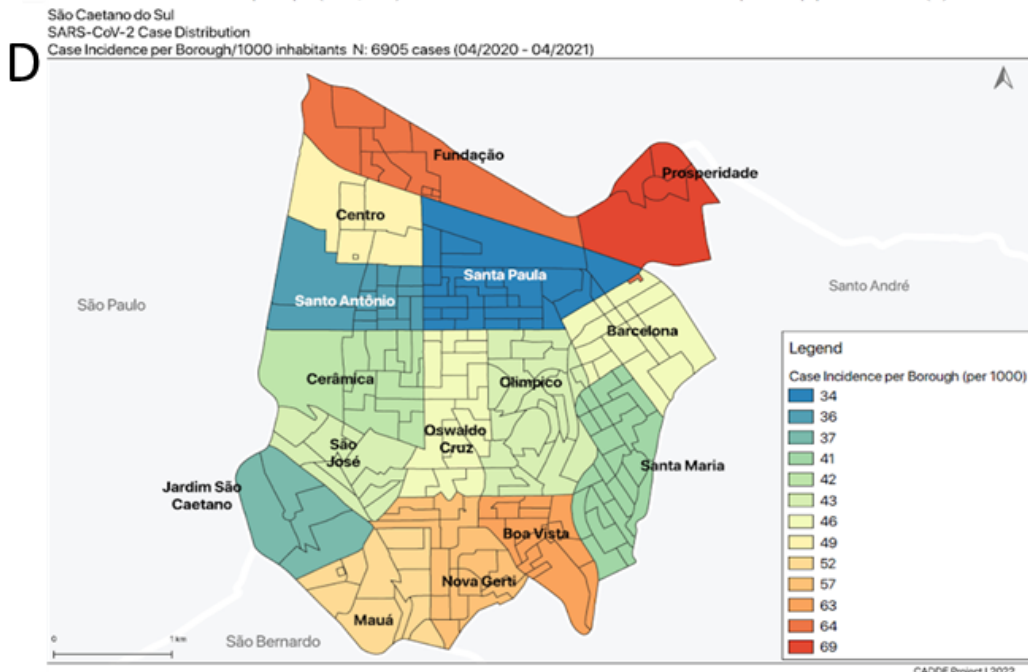
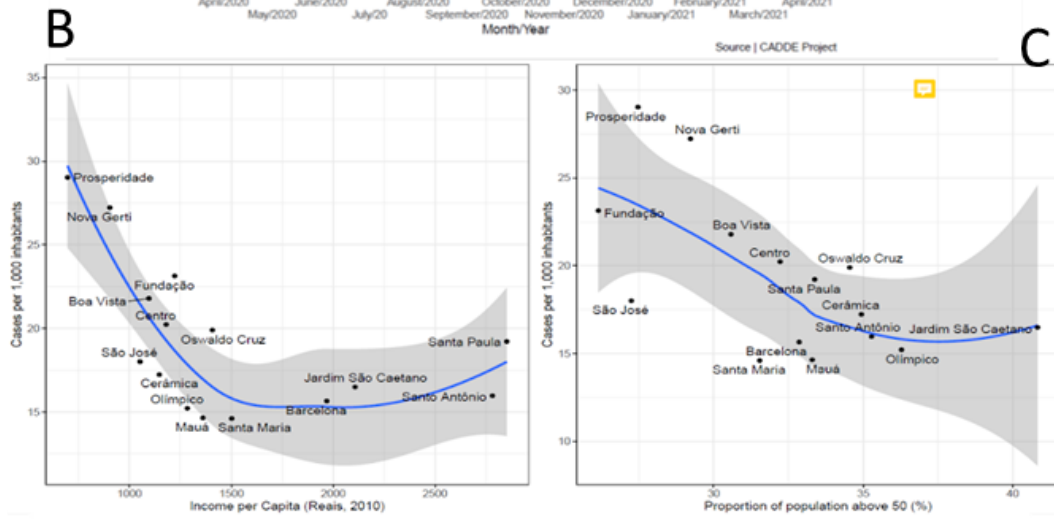
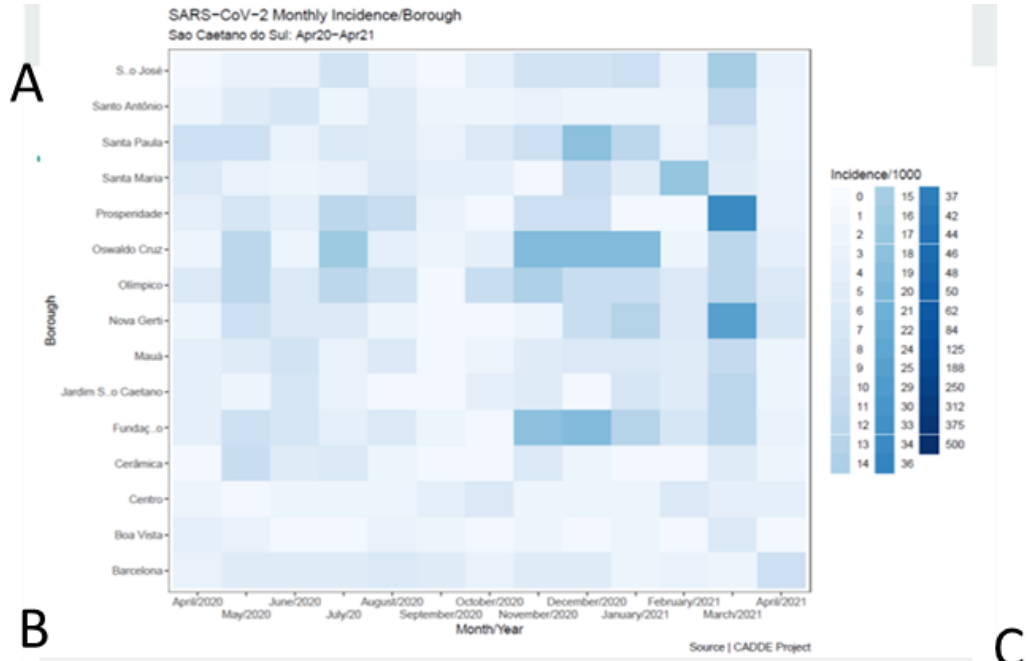
Legenda. Gráfico de correlação de Pearson demonstrando um conjunto de dados fortemente correlacionado com casos confirmados de SARS CoV-2 PCR+ espacialmente (A) e temporalmente (B) (correlação de Pearson = 0,87 e 0,45,

respectivamente). A linha é a função linear (sem interceptação) que melhor ajusta os pontos de dados no sentido do quadrado médio.

4.4 Análises epidemiológicas

Além disso, investigamos a incidência mensal de casos de COVID-19 por bairro na cidade de SCS (**Figura 10A**), considerando renda per capita e idade (**Figuras 11B e 11C**). Os resultados mostraram que a incidência variou de 34 a 69 casos por 1.000 habitantes, os bairros com menor renda per capita e com população mais jovem tiveram maior incidência de casos de COVID-19. O mapa cumulativo do total de casos por bairro é apresentado na **Figura 10D**.

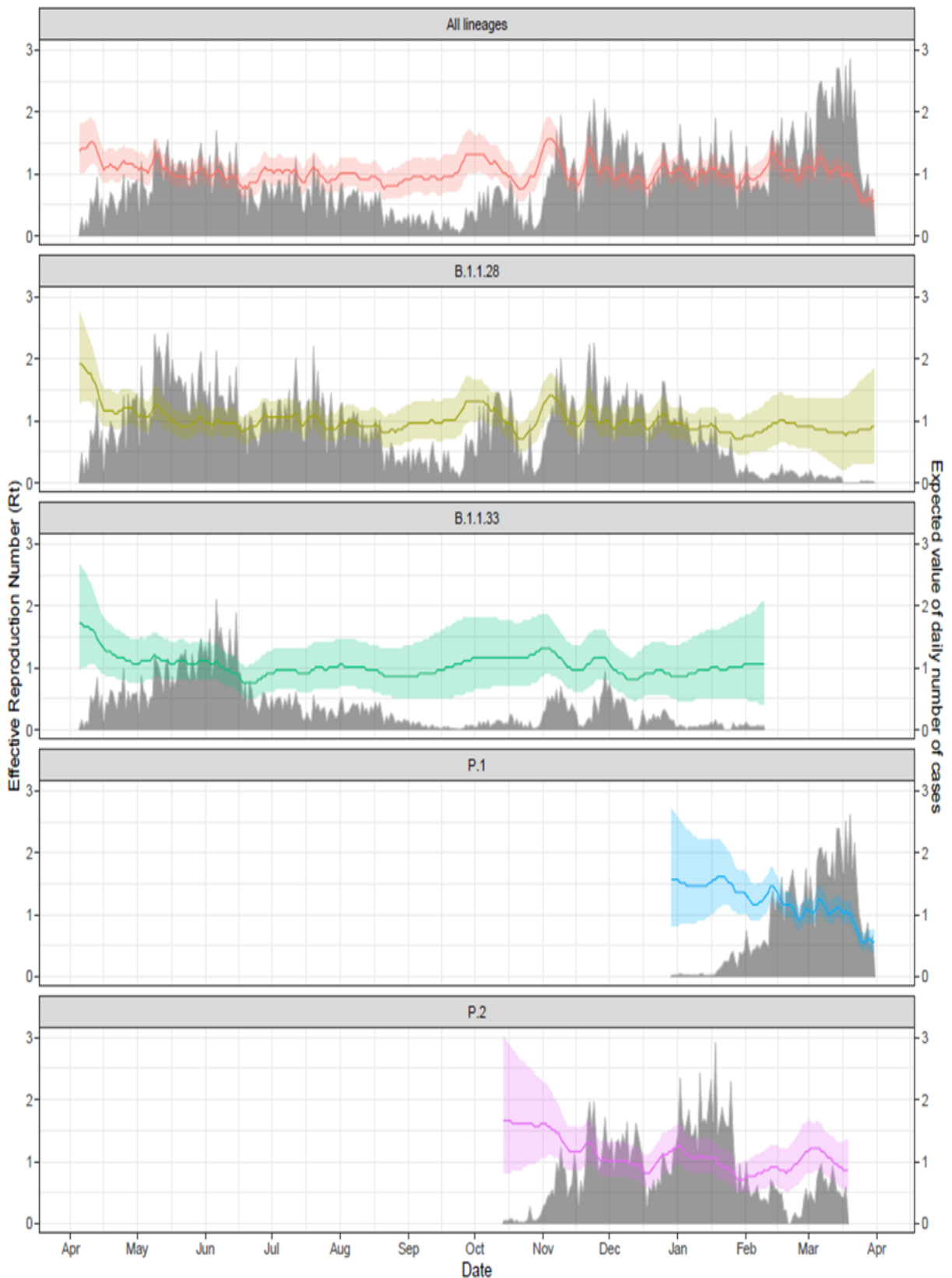
Figura 10. Mapas de incidência mensal e cumulativos em São Caetano do Sul.



Legenda. Mapa de calor demonstrando a incidência mensal de SARS-CoV-2 por bairro na cidade de São Caetano do Sul entre abril de 2020 - abril de 2021. Os bairros estão sombreados de acordo com o número de casos amostrados. Gráfico de regressão mostrando maior incidência (1000/habitante) de casos de COVID-19 em bairros mais pobres (B) e com população (C) mais jovem (dados do IBGE/2010). **D)** Mapa cumulativo de incidência de casos (n=6905) por 1000 habitantes. Os bairros são sombreados de acordo com o número de casos amostrados.

O valor esperado do número diário de casos e o número efetivo de reprodução (R_t) foram calculados para as principais cepas que circularam no período estudado (B.1.1.28, B.1.1.33, Zeta/VOI) e Gama/VOC). Durante a primeira e segunda fases do estudo, o R_t oscilou entre 1 e 1,5 para ambas as cepas ancestrais (B.1.1.28 e B.1.1.33). Em novembro de 2020, foi observado um R_t alto ($>1,5$) para a linhagem B.1.1.28, coincidindo com o surgimento da linhagem Zeta (P.2). Nos dois primeiros meses da epidemia de Gama, observou-se uma oscilação do R_t entre 1 e 1,5. (**Gráfico 3**).

Gráfico 3. Número efetivo de reprodução (R_t) por linhagem

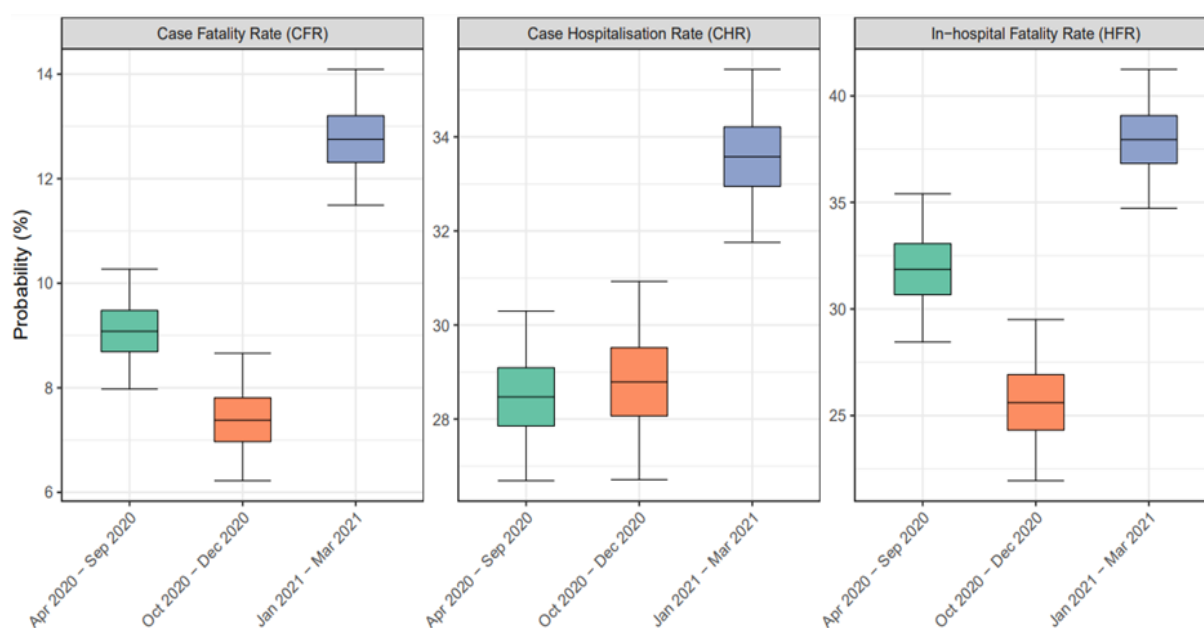


Legenda. As linhas mostram o número efetivo de reprodução (R_t) (eixo y à esquerda) estimado para cada linhagem e as faixas indicam seus intervalos de 95% de credibilidade. As barras cinzas representam o valor esperado do número diário estimado de casos (eixo y à direita) causados por cada uma das principais linhagens

(B.1.1.28, B.1.1.33, P.1 e P2) circulando entre abril de 2020 e abril de 2021.

Análises descritivas da taxa de letalidade (Case Fatality Rate - CFR), taxa de hospitalização (Case Hospitalisation Rate - CHR) e taxa de letalidade intra-hospitalar (In-hospital Fatality Rate - HFR) foram realizadas cruzando os dados do SIVEP-GRIPE (Sistema de Informação de Vigilância Epidemiológica da Influenza) e dados da Plataforma SC. Todos esses indicadores foram mais elevados durante o terceiro período, quando a linhagem Gama/VOC foi prevalente (**Gráfico 4**).

Gráfico 4. Análises descritivas de taxa de hospitalização, taxa de letalidade intra-hospitalar e taxa de hospitalar

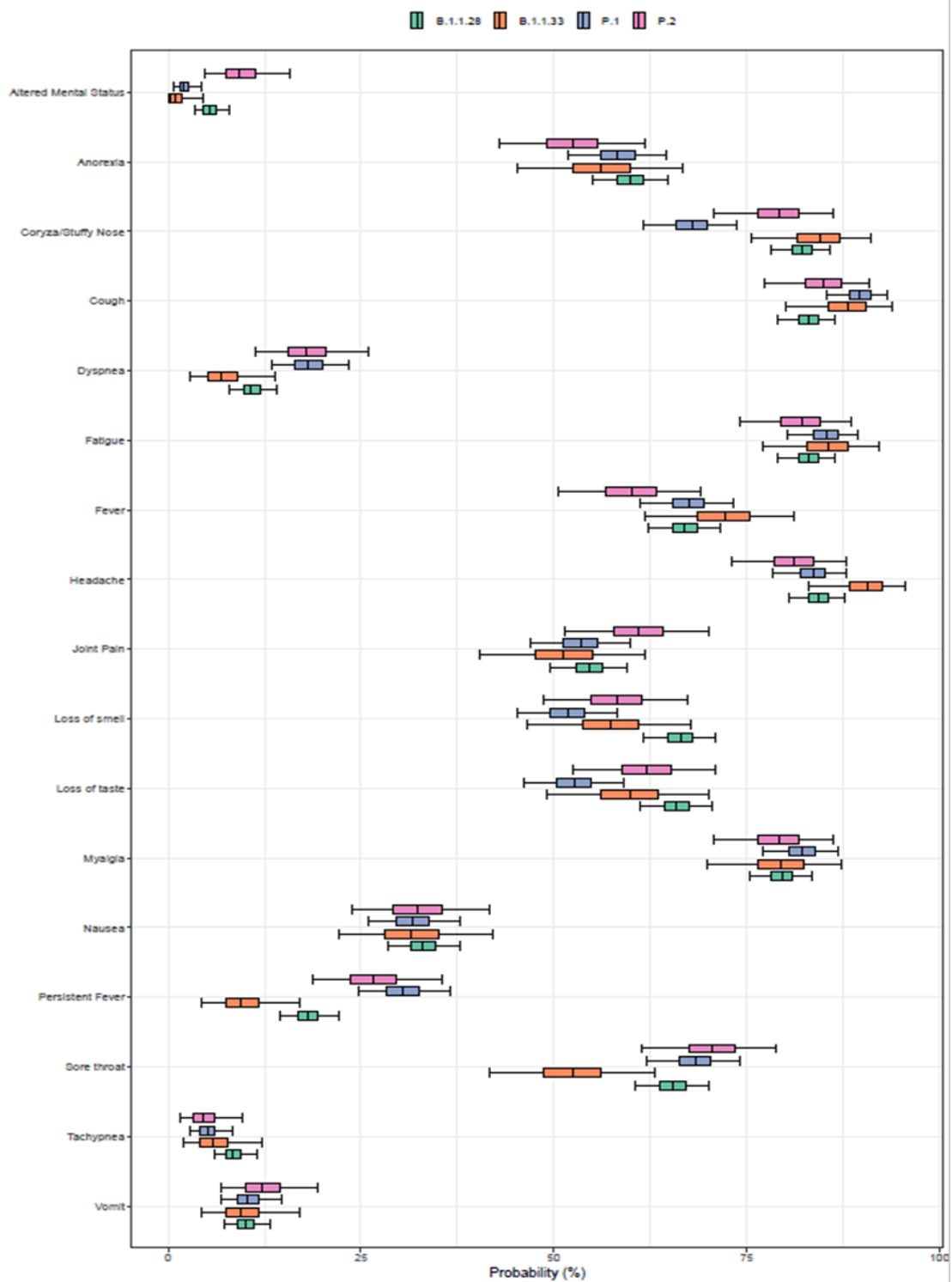


Legenda. Análises descritivas foram realizadas para estabelecer a taxa de letalidade (CFR), taxa de hospitalização (CHR) e taxa de letalidade intra-hospitalar (HFR) desagregadas por período de estudo.

Traçamos a proporção de cada sintoma específico de acordo com as quatro linhagens mais comuns. Avaliamos os 879 casos sequenciados (**Gráfico 5**) e os 6.095 casos positivos (**Gráfico 6**) após inserir a variante mais provável de acordo com a semana do início dos sintomas. Ambos os gráficos

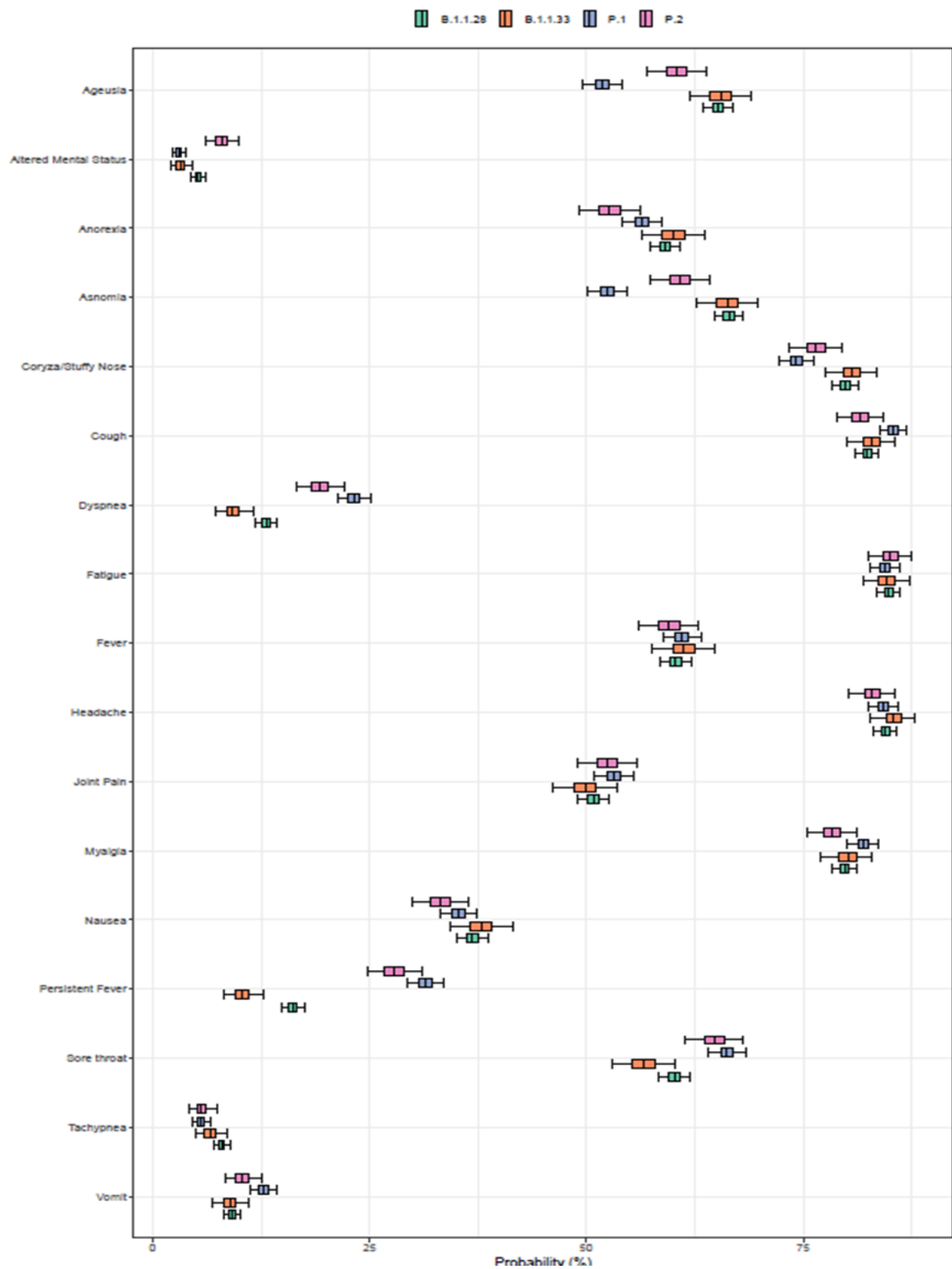
mostram que os sintomas relacionados à doença mais grave, como dispnéia e febre prolongada, foram aproximadamente duas vezes mais comuns entre as linhagens Gama/VOC e Zeta/VOI em comparação com as variantes ancestrais (B1.1.28 e B1.1.33). Ageusia e anosmia foram menos comuns na linhagem Gama e sintomas neurológicos foram mais comuns na linhagem Zeta. Também demonstramos que a gravidade medida pela dispneia foi maior durante a segunda e terceira fase para todas as faixas de idade adulta (**Gráfico 7**).

Gráfico 5. Probabilidade do sintoma ocorrer de acordo com a linhagem atribuída à amostra utilizando todos os casos sequenciados.



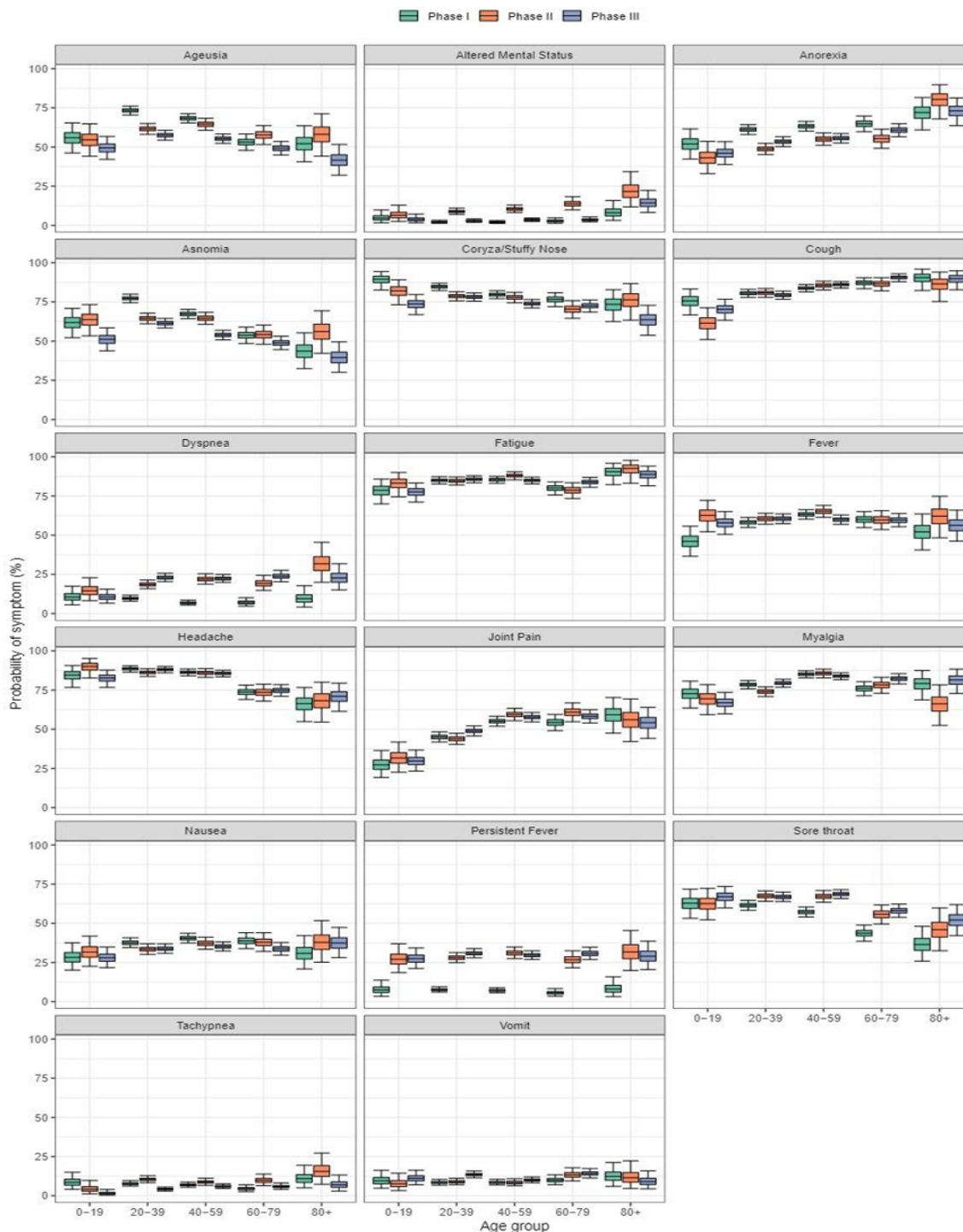
Legenda. Os quantis da distribuição *Beta* são desenhados para estimar a mediana e os intervalos de 95% de confiança.

Gráfico 6. Probabilidade do sintoma ocorrer de acordo com a linhagem atribuída à amostra utilizando todos os positivos para COVID-19.



Legenda. Modelo imputado, onde consideramos todas as amostras com PCR+ para SARS-CoV-2. Os quantis da distribuição *Beta* são desenhados para estimar a mediana e os intervalos de 95% de confiança.

Gráfico 7. Probabilidade de ocorrência de determinado sintoma de acordo com a faixa etária e fase do estudo (I, II e III).



Legenda. As caixas mostram as medianas, 50% dos intervalos de confiança e os bigodes indicam 95% dos intervalos de confiança. Os sinais de alarme para severidade são: febre/febre persistente, dispnéia/dificuldade para respirar, taquipneia e status mental alterado

5. DISCUSSÃO

Apresentamos a análise dos dados obtidos a partir de uma plataforma web criada para atender casos de COVID-19 no município brasileiro de São Caetano do Sul (152) que representam 63% de todos os casos notificados na cidade. Além dos dados epidemiológicos e clínicos, amostras positivas de PCR puderam ser recuperadas para sequenciamento, permitindo uma descrição detalhada da evolução da epidemia no município e a comparação da apresentação da doença entre as diferentes linhagens.

Apesar dos esforços de sequenciamento, não há conhecimento preciso sobre a frequência de distribuição das linhagens SARS-CoV-2 no Brasil, principalmente devido às lacunas amostrais nos estratos espaço-temporais e ao baixo número de genomas de regiões, como norte e nordeste do Brasil. Montar um conjunto de dados que seja representativo de todas as regiões brasileiras é uma tarefa desafiadora. Nessa análise do relógio molecular, data a raiz da árvore global SARS-CoV-2 em Dezembro de 2019 (95% da densidade posterior mais alta (HPD): 8 de Dezembro de 2019 a 28 de Dezembro de 2019), consistente com descobertas anteriores. Nossa análise também suporta a ocorrência de múltiplas introduções do vírus no Brasil no início de 2020, que culminaram no estabelecimento de linhagens brasileiras ao longo do primeiro semestre daquele ano, conforme relatado anteriormente (91). As sequências de São Caetano do Sul agruparam-se com múltiplos clados espalhados na árvore, denotando múltiplas introduções, independentes de diversas linhagens (B.1.1.28 e P1, principalmente), inclusive de outros estados brasileiros em todas as fases do estudo, que podem ser explicada devido a mobilidade urbana (a cidade de São Caetano do Sul é intensamente conurbada com o município de São Paulo e a região metropolitana de São Paulo, sem limites físicos, possui linha de trens e está próximo ao aeroporto Internacional de São Paulo/Guarulhos e ao aeroporto de Congonhas).

Pudemos mostrar que a apresentação da doença difere entre as linhagens: as Figuras **1S** e **2S** e a **Tabela 2S** mostram que Gama e Zeta foram significativamente mais propensos a apresentar sintomas mais graves como dispneia (1,81 e 1,79 vezes) e febre prolongada (1,85 e 1,61 vezes) quando

comparado com as linhagens ancestrais. Anosmia e ageusia foram menos frequentes entre os gamma VOC, enquanto os sintomas mentais relacionados foram mais presentes entre os 61 casos infectados com Zeta VOI.

O Brasil passou por uma das epidemias mais significativas do mundo, com importantes diferenças regionais. A epidemia na cidade de São Caetano do Sul foi semelhante à da região metropolitana de São Paulo com uma primeira onda prolongada e em forma de platô, seguida por uma segunda onda mais proeminente impulsionada inicialmente pela linhagem Zeta, que foi superada pela Gama VOC em Fevereiro de 2021. Estima-se que até dezembro de 2020 26.6% da população da cidade de São Paulo havia sido infectada (176).

A linhagem Gamma foi detectada pela primeira vez em Manaus, uma região que alcançou uma alta taxa de ataque durante a primeira onda (92,120) e rapidamente se expandiu para todas as cidades brasileiras causando um aumento dramático de mortes por COVID-19. No trabalho original que descreve a linhagem Gama, sugerimos que a P.1 provavelmente causou 1,2 a 1,9 vezes mais mortes (92). No entanto, quando (177) analisaram dados de pacientes hospitalizados do Brasil, concluíram que as taxas de mortalidade intra-hospitalar estavam associadas principalmente a desigualdades geográficas e escassez de capacidade de assistência médica e não à linhagem Gama. Outros estudos, no entanto, apontam para uma maior patogenicidade do Gama/VOC, Banho et al. (178) avaliaram a taxa de casos graves e não graves na cidade de São José dos Campos no Brasil e detectaram uma maior proporção de casos graves casos, especialmente entre a população mais jovem após a introdução da linhagem Gama.

Ao avaliar a taxa de ataque em Manaus durante a primeira e a segunda onda, Prete et al. (175) concluíram que a taxa de infecção para hospitalização e infecção para fatalidade foi maior durante o período da linhagem Gama. Nosso estudo também sugere um desfecho mais grave para a infecção por Gama/VOC, pois os pacientes infectados por essa variante relataram com mais frequência dispneia e febre prolongada. Além disso, durante o período em que o Gama foi responsável pela maioria dos casos, a taxa de letalidade (CFR), a taxa de hospitalização (CHR) e a taxa de letalidade intra-hospitalar (HFR) foram maiores. Apesar dos sintomas mais graves associados à infecção por

Zeta/VOI, não houve aumento significativo na taxa de internação no período. Anosmia e ageusia foram detectadas com menos frequência em pacientes infectados por Gama/VOC e isso também foi relatado por Luna-Muschi et al. (179) em uma coorte de profissionais de saúde do maior hospital de São Paulo. Curiosamente, a infecção com Zeta/VOI apresentou mais sintomas neurológicos e isso não foi relatado anteriormente.

A introdução de VOCs internacionais em território brasileiro é um capítulo importante da história da evolução e disseminação de SARS-CoV-2. Os dois primeiros casos da linhagem Alfa (86) foram identificados no final do ano de 2020 através da vigilância genômica. Utilizando-se de análises filogenéticas, constatou-se que eram duas introduções diferentes. Em nosso estudo, identificamos 7 sequências de linhagem Alfa/VOC (B.1.1.7), porém não conseguimos relacionar vínculos epidemiológicos.

Estudos com amostragem representativa e oportunista de uma região é fundamental para os objetivos de vigilância genômica. A utilização de um esquema de amostragem adequada, independentemente dos metadados da amostra, juntamente com a integração de outras fontes de informação padronizadas, como dados epidemiológicos oficiais e parâmetros mais sensíveis utilizados em análises computacionais, contribuem para uma interpretação mais confiável, reduzindo problemas de subamostragem e vieses e aumentando a possibilidade de identificar linhagens emergentes (180); (121,181). Essas estratégias serão cada vez mais necessárias agora que é amplamente reconhecido que os estudos de todo o genoma serão mais acessíveis e usados rotineiramente (182,183). Nosso banco de dados é único e conseguimos incluir um grande número de amostras, de todas as regiões da cidade. A Plataforma São Caetano pôde contribuir em momento delicado na história da pandemia, no início da quarentena, em um estágio em que a recomendação era de permanecer em casa e evitar deslocamentos necessários. Realizou atendimentos médicos através das plataformas digitais, realizando triagem de casos potencialmente graves e encaminhando-os para atendimento médico de urgência e principalmente, disponibilizando recursos técnicos e disponibilizou testes de diagnóstico de forma menos burocrática à população, principalmente em um período, que insumos médicos e

laboratoriais eram escassos (91) e além disso, contribuiu para a diminuição da subnotificação de casos de COVID-19 no período. Dessa forma, nosso estudo mostra que é possível obter insights sólidos sobre a transmissão do COVID-19, quando a diversidade genômica do vírus é combinada e interpretada em conjunto com dados epidemiológicos detalhados dos casos.

6. CONCLUSÃO

As informações obtidas neste estudo podem ser utilizadas para nortear políticas públicas e intervenções em saúde pública na cidade de São Caetano do Sul e em outras localidades com características semelhantes. Adicionalmente, a coleta de dados por meio de plataformas digitais e dados fornecidos pelo sequenciamento genômico, são cruciais para a vigilância epidemiológica e controle da disseminação de doenças infecciosas, permitindo uma coleta de dados mais ágil e eficiente.

7. MATERIAL SUPLEMENTAR

Gráfico 1S. Múltiplas comparações entre os principais sintomas apresentados (n=17) e as linhagens identificadas (n=4), corrigidas por Bonferroni, considerando o p-value de 0,05.

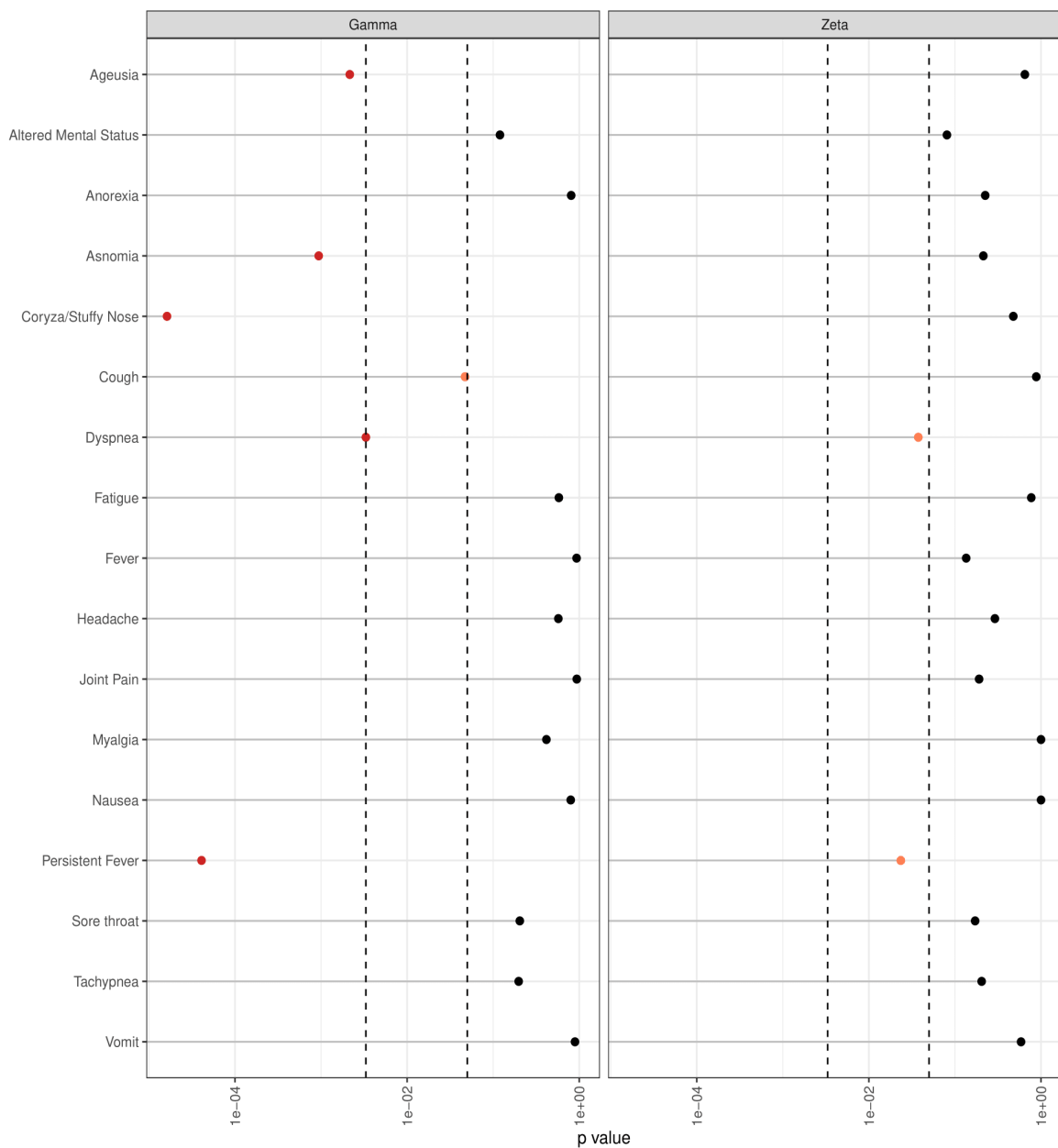


Tabela 1S - Risco relativo de ocorrência de determinado sintoma tendo como referência a prevalência desse sintoma para as linhagens B.1.1.28 e B.1.1.33.

Sintomas	Gamma vs B.1.1.28 e B.1.1.33		Zeta vs B.1.1.28 e B.1.1.33	
	Risco Relativo (95% CrI)	p-value	Risco Relativo (95% CrI)	p-value
Ageusia	0.81 (0.70 - 0.93)	0,0022	0.96 (0.80 - 1.11)	0,6493
Status mental alterado	0.46 (0.15 - 1.12)	0,1198	2.09 (0.95 - 4.17)	0,0807
Anorexia	0.98 (0.86 - 1.11)	0,8057	0.88 (0.71 - 1.06)	0,2248
Anosmia	0.80 (0.69 - 0.91)	0,0009	0.90 (0.74 - 1.06)	0,2138
Coriza/nariz entupido	0.82 (0.74 - 0.90)	1,6208	0.96 (0.85 - 1.05)	0,4765
Tosse	1.07 (1.00 - 1.13)	0,0471	1.01 (0.91 - 1.10)	0,8809
Dispneia	1.81 (1.22 - 2.66)	0,0033	1.79 (1.07 - 2.83)	0,0375
Fadiga	1.02 (0.95 - 1.09)	0,5794	0.98 (0.88 - 1.07)	0,7705
Febre	0.99 (0.89 - 1.10)	0,9312	0.89 (0.74 - 1.03)	0,1350
Dor de Cabeça	0.98 (0.91 - 1.04)	0,5716	0.95 (0.85 - 1.04)	0,2913
Dor na articulação	0.99 (0.85 - 1.14)	0,9354	1.13 (0.94 - 1.33)	0,1908
Mialgia	1.03 (0.95 - 1.11)	0,4153	0.99 (0.88 - 1.10)	1,0000
Náusea	0.97 (0.77 - 1.22)	0,7958	0.99 (0.71 - 1.32)	1,0000
Febre Persistente	1.85 (1.40 - 2.45)	4,0836	1.61 (1.08 - 2.35)	0,0235
Dor de garganta	1.08 (0.96 - 1.21)	0,2038	1.12 (0.96 - 1.28)	0,1716
Taquipneia	0.64 (0.33 - 1.18)	0,1977	0.57 (0.20 - 1.30)	0,2043
Vômito	1.05 (0.64 - 1.66)	0,8922	1.23 (0.65 - 2.15)	0,5863

Tabela 2S. Lista com os IDs das sequências geradas submetidas à Plataforma online GISAID.

EPI_ISL_17153240, EPI_ISL_17153251, EPI_ISL_17153250, EPI_ISL_17153242,
EPI_ISL_17153253, EPI_ISL_17153241, EPI_ISL_17153252, EPI_ISL_17153244,
EPI_ISL_17153255, EPI_ISL_17153243, EPI_ISL_17153254, EPI_ISL_17153246,
EPI_ISL_17153245, EPI_ISL_17153256, EPI_ISL_17153237, EPI_ISL_17153248,
EPI_ISL_17153247, EPI_ISL_17153239, EPI_ISL_17153238, EPI_ISL_17153249,
EPI_ISL_17127604, EPI_ISL_17127605, EPI_ISL_17127606, EPI_ISL_17127607,
EPI_ISL_17127608, EPI_ISL_17127609, EPI_ISL_17127585, EPI_ISL_17127586,
EPI_ISL_17127587, EPI_ISL_17127588, EPI_ISL_17127589, EPI_ISL_17127600,
EPI_ISL_17127601, EPI_ISL_17127602, EPI_ISL_17127603, EPI_ISL_17127590,
EPI_ISL_17127591, EPI_ISL_17127570, EPI_ISL_17127592, EPI_ISL_17127571,
EPI_ISL_17127593, EPI_ISL_17127572, EPI_ISL_17127594, EPI_ISL_17127573,
EPI_ISL_17127595, EPI_ISL_17127574, EPI_ISL_17127596, EPI_ISL_17127575,
EPI_ISL_17127597, EPI_ISL_17127576, EPI_ISL_17127598, EPI_ISL_17127577,
EPI_ISL_17127599, EPI_ISL_17127610, EPI_ISL_17127578, EPI_ISL_17127611,
EPI_ISL_17127579, EPI_ISL_17127612, EPI_ISL_17127613, EPI_ISL_17127614,
EPI_ISL_17127580, EPI_ISL_17127581, EPI_ISL_17127582, EPI_ISL_17127583,
EPI_ISL_17127584, EPI_ISL_17050139, EPI_ISL_17050376, EPI_ISL_17050134,
EPI_ISL_17050375, EPI_ISL_17050133, EPI_ISL_17050374, EPI_ISL_17050132,
EPI_ISL_17050373, EPI_ISL_17050131, EPI_ISL_17050138, EPI_ISL_17050137,
EPI_ISL_17050379, EPI_ISL_17050378, EPI_ISL_17050136, EPI_ISL_17050377,
EPI_ISL_17050135, EPI_ISL_17050372, EPI_ISL_17050130, EPI_ISL_17050371,
EPI_ISL_17050370, EPI_ISL_17050129, EPI_ISL_17050128, EPI_ISL_17050365,
EPI_ISL_17050123, EPI_ISL_17050364, EPI_ISL_17050122, EPI_ISL_17050363,
EPI_ISL_17050121, EPI_ISL_17050362, EPI_ISL_17050120, EPI_ISL_17050127,
EPI_ISL_17050369, EPI_ISL_17050368, EPI_ISL_17050126, EPI_ISL_17050367,
EPI_ISL_17050125, EPI_ISL_17050366, EPI_ISL_17050124, EPI_ISL_17050361,
EPI_ISL_17050360, EPI_ISL_17050156, EPI_ISL_17050398, EPI_ISL_17050155,
EPI_ISL_17050397, EPI_ISL_17050154, EPI_ISL_17050396, EPI_ISL_17050153,
EPI_ISL_17050395, EPI_ISL_17050159, EPI_ISL_17050158, EPI_ISL_17050157,
EPI_ISL_17050399, EPI_ISL_17050390, EPI_ISL_17050152, EPI_ISL_17050394,
EPI_ISL_17050151, EPI_ISL_17050393, EPI_ISL_17050150, EPI_ISL_17050392,
EPI_ISL_17050391, EPI_ISL_17050145, EPI_ISL_17050387, EPI_ISL_17050144,
EPI_ISL_17050386, EPI_ISL_17050143, EPI_ISL_17050385, EPI_ISL_17050142,
EPI_ISL_17050384, EPI_ISL_17050149, EPI_ISL_17050148, EPI_ISL_17050147,

EPI_ISL_17050389, EPI_ISL_17050146, EPI_ISL_17050388, EPI_ISL_17050141,
EPI_ISL_17050383, EPI_ISL_17050140, EPI_ISL_17050382, EPI_ISL_17050381,
EPI_ISL_17050380, EPI_ISL_17050178, EPI_ISL_17050177, EPI_ISL_17050176,
EPI_ISL_17050175, EPI_ISL_17050179, EPI_ISL_17050170, EPI_ISL_17050174,
EPI_ISL_17050173, EPI_ISL_17050172, EPI_ISL_17050171, EPI_ISL_17050167,
EPI_ISL_17050166, EPI_ISL_17050165, EPI_ISL_17050164, EPI_ISL_17050169,
EPI_ISL_17050168, EPI_ISL_17050163, EPI_ISL_17050162, EPI_ISL_17050161,
EPI_ISL_17050160, EPI_ISL_17050199, EPI_ISL_17050198, EPI_ISL_17050197,
EPI_ISL_17050192, EPI_ISL_17050191, EPI_ISL_17050190, EPI_ISL_17050196,
EPI_ISL_17050195, EPI_ISL_17050194, EPI_ISL_17050193, EPI_ISL_17050189,
EPI_ISL_17050188, EPI_ISL_17050187, EPI_ISL_17050186, EPI_ISL_17050181,
EPI_ISL_17050180, EPI_ISL_17050185, EPI_ISL_17050184, EPI_ISL_17050183,
EPI_ISL_17050182, EPI_ISL_17050538, EPI_ISL_17050537, EPI_ISL_17050536,
EPI_ISL_17050535, EPI_ISL_17050539, EPI_ISL_17050530, EPI_ISL_17050534,
EPI_ISL_17050533, EPI_ISL_17050532, EPI_ISL_17050531, EPI_ISL_17050527,
EPI_ISL_17050526, EPI_ISL_17050525, EPI_ISL_17050524, EPI_ISL_17050529,
EPI_ISL_17050528, EPI_ISL_17050523, EPI_ISL_17050522, EPI_ISL_17050521,
EPI_ISL_17050520, EPI_ISL_17050318, EPI_ISL_17050317, EPI_ISL_17050559,
EPI_ISL_17050558, EPI_ISL_17050316, EPI_ISL_17050557, EPI_ISL_17050315,
EPI_ISL_17050319, EPI_ISL_17050552, EPI_ISL_17050310, EPI_ISL_17050551,
EPI_ISL_17050550, EPI_ISL_17050556, EPI_ISL_17050314, EPI_ISL_17050555,
EPI_ISL_17050313, EPI_ISL_17050554, EPI_ISL_17050312, EPI_ISL_17050553,
EPI_ISL_17050311, EPI_ISL_17050307, EPI_ISL_17050549, EPI_ISL_17050548,
EPI_ISL_17050306, EPI_ISL_17050547, EPI_ISL_17050305, EPI_ISL_17050546,
EPI_ISL_17050304, EPI_ISL_17050309, EPI_ISL_17050308, EPI_ISL_17050541,
EPI_ISL_17050540, EPI_ISL_17050545, EPI_ISL_17050303, EPI_ISL_17050544,
EPI_ISL_17050302, EPI_ISL_17050543, EPI_ISL_17050301, EPI_ISL_17050542,
EPI_ISL_17050300, EPI_ISL_17050339, EPI_ISL_17050338, EPI_ISL_17050337,
EPI_ISL_17050579, EPI_ISL_17050574, EPI_ISL_17050332, EPI_ISL_17050573,
EPI_ISL_17050331, EPI_ISL_17050572, EPI_ISL_17050330, EPI_ISL_17050571,
EPI_ISL_17050578, EPI_ISL_17050336, EPI_ISL_17050577, EPI_ISL_17050335,
EPI_ISL_17050576, EPI_ISL_17050334, EPI_ISL_17050575, EPI_ISL_17050333,
EPI_ISL_17050570, EPI_ISL_17050329, EPI_ISL_17050328, EPI_ISL_17050327,
EPI_ISL_17050569, EPI_ISL_17050568, EPI_ISL_17050326, EPI_ISL_17050563,
EPI_ISL_17050321, EPI_ISL_17050562, EPI_ISL_17050320, EPI_ISL_17050561,
EPI_ISL_17050560, EPI_ISL_17050567, EPI_ISL_17050325, EPI_ISL_17050566,
EPI_ISL_17050324, EPI_ISL_17050565, EPI_ISL_17050323, EPI_ISL_17050564,

EPI_ISL_17050322, EPI_ISL_17050119, EPI_ISL_17050118, EPI_ISL_17050117,
EPI_ISL_17050359, EPI_ISL_17050354, EPI_ISL_17050112, EPI_ISL_17050596,
EPI_ISL_17050353, EPI_ISL_17050111, EPI_ISL_17050595, EPI_ISL_17050352,
EPI_ISL_17050110, EPI_ISL_17050594, EPI_ISL_17050351, EPI_ISL_17050593,
EPI_ISL_17050358, EPI_ISL_17050116, EPI_ISL_17050357, EPI_ISL_17050115,
EPI_ISL_17050599, EPI_ISL_17050356, EPI_ISL_17050114, EPI_ISL_17050598,
EPI_ISL_17050355, EPI_ISL_17050113, EPI_ISL_17050597, EPI_ISL_17050350,
EPI_ISL_17050592, EPI_ISL_17050591, EPI_ISL_17050590, EPI_ISL_17050109,
EPI_ISL_17050108, EPI_ISL_17050107, EPI_ISL_17050349, EPI_ISL_17050348,
EPI_ISL_17050106, EPI_ISL_17050343, EPI_ISL_17050101, EPI_ISL_17050585,
EPI_ISL_17050342, EPI_ISL_17050100, EPI_ISL_17050584, EPI_ISL_17050341,
EPI_ISL_17050583, EPI_ISL_17050340, EPI_ISL_17050582, EPI_ISL_17050347,
EPI_ISL_17050105, EPI_ISL_17050589, EPI_ISL_17050346, EPI_ISL_17050104,
EPI_ISL_17050588, EPI_ISL_17050345, EPI_ISL_17050103, EPI_ISL_17050587,
EPI_ISL_17050344, EPI_ISL_17050102, EPI_ISL_17050586, EPI_ISL_17050581,
EPI_ISL_17050580, EPI_ISL_17049947, EPI_ISL_17049946, EPI_ISL_17049949,
EPI_ISL_17049948, EPI_ISL_17050714, EPI_ISL_17050713, EPI_ISL_17050712,
EPI_ISL_17050711, EPI_ISL_17050718, EPI_ISL_17050717, EPI_ISL_17050716,
EPI_ISL_17050715, EPI_ISL_17049961, EPI_ISL_17049960, EPI_ISL_17049963,
EPI_ISL_17050710, EPI_ISL_17049962, EPI_ISL_17049965, EPI_ISL_17049964,
EPI_ISL_17049967, EPI_ISL_17049966, EPI_ISL_17049969, EPI_ISL_17049968,
EPI_ISL_17050709, EPI_ISL_17050708, EPI_ISL_17050703, EPI_ISL_17050702,
EPI_ISL_17050701, EPI_ISL_17050700, EPI_ISL_17050707, EPI_ISL_17050706,
EPI_ISL_17050705, EPI_ISL_17050704, EPI_ISL_17049950, EPI_ISL_17049952,
EPI_ISL_17049951, EPI_ISL_17049954, EPI_ISL_17049953, EPI_ISL_17049956,
EPI_ISL_17049955, EPI_ISL_17049958, EPI_ISL_17049957, EPI_ISL_17049959,
EPI_ISL_17049990, EPI_ISL_17049981, EPI_ISL_17049980, EPI_ISL_17049983,
EPI_ISL_17049982, EPI_ISL_17049985, EPI_ISL_17049984, EPI_ISL_17049987,
EPI_ISL_17049986, EPI_ISL_17049989, EPI_ISL_17049988, EPI_ISL_17049970,
EPI_ISL_17049972, EPI_ISL_17049971, EPI_ISL_17049974, EPI_ISL_17049973,
EPI_ISL_17049976, EPI_ISL_17049975, EPI_ISL_17049978, EPI_ISL_17049977,
EPI_ISL_17049979, EPI_ISL_17050516, EPI_ISL_17050515, EPI_ISL_17050514,
EPI_ISL_17050513, EPI_ISL_17050519, EPI_ISL_17050518, EPI_ISL_17050517,
EPI_ISL_17050512, EPI_ISL_17050511, EPI_ISL_17050510, EPI_ISL_17050505,
EPI_ISL_17050504, EPI_ISL_17050503, EPI_ISL_17050502, EPI_ISL_17050509,
EPI_ISL_17050508, EPI_ISL_17050507, EPI_ISL_17050506, EPI_ISL_17049992,
EPI_ISL_17049991, EPI_ISL_17049994, EPI_ISL_17049993, EPI_ISL_17050501,

EPI_ISL_17049996, EPI_ISL_17050500, EPI_ISL_17049995, EPI_ISL_17049998,
EPI_ISL_17049997, EPI_ISL_17049999, EPI_ISL_17050099, EPI_ISL_17050098,
EPI_ISL_17050093, EPI_ISL_17050092, EPI_ISL_17050091, EPI_ISL_17050090,
EPI_ISL_17050097, EPI_ISL_17050096, EPI_ISL_17050095, EPI_ISL_17050094,
EPI_ISL_17050089, EPI_ISL_17050088, EPI_ISL_17050087, EPI_ISL_17050082,
EPI_ISL_17050081, EPI_ISL_17050080, EPI_ISL_17050086, EPI_ISL_17050085,
EPI_ISL_17050084, EPI_ISL_17050083, EPI_ISL_17050019, EPI_ISL_17050018,
EPI_ISL_17050255, EPI_ISL_17050013, EPI_ISL_17050497, EPI_ISL_17050254,
EPI_ISL_17050012, EPI_ISL_17050496, EPI_ISL_17050253, EPI_ISL_17050011,
EPI_ISL_17050495, EPI_ISL_17050252, EPI_ISL_17050010, EPI_ISL_17050494,
EPI_ISL_17050017, EPI_ISL_17050259, EPI_ISL_17050258, EPI_ISL_17050016,
EPI_ISL_17050257, EPI_ISL_17050015, EPI_ISL_17050499, EPI_ISL_17050256,
EPI_ISL_17050014, EPI_ISL_17050498, EPI_ISL_17050251, EPI_ISL_17050493,
EPI_ISL_17050250, EPI_ISL_17050492, EPI_ISL_17050491, EPI_ISL_17050490,
EPI_ISL_17050009, EPI_ISL_17050008, EPI_ISL_17050007, EPI_ISL_17050249,
EPI_ISL_17050244, EPI_ISL_17050002, EPI_ISL_17050486, EPI_ISL_17050243,
EPI_ISL_17050001, EPI_ISL_17050485, EPI_ISL_17050242, EPI_ISL_17050000,
EPI_ISL_17050484, EPI_ISL_17050241, EPI_ISL_17050483, EPI_ISL_17050248,
EPI_ISL_17050006, EPI_ISL_17050247, EPI_ISL_17050005, EPI_ISL_17050489,
EPI_ISL_17050246, EPI_ISL_17050004, EPI_ISL_17050488, EPI_ISL_17050245,
EPI_ISL_17050003, EPI_ISL_17050487, EPI_ISL_17050240, EPI_ISL_17050482,
EPI_ISL_17050481, EPI_ISL_17050480, EPI_ISL_17050277, EPI_ISL_17050035,
EPI_ISL_17050276, EPI_ISL_17050034, EPI_ISL_17050275, EPI_ISL_17050033,
EPI_ISL_17050274, EPI_ISL_17050032, EPI_ISL_17050039, EPI_ISL_17050038,
EPI_ISL_17050037, EPI_ISL_17050279, EPI_ISL_17050278, EPI_ISL_17050036,
EPI_ISL_17050273, EPI_ISL_17050031, EPI_ISL_17050272, EPI_ISL_17050030,
EPI_ISL_17050271, EPI_ISL_17050270, EPI_ISL_17050029, EPI_ISL_17050266,
EPI_ISL_17050024, EPI_ISL_17050265, EPI_ISL_17050023, EPI_ISL_17050264,
EPI_ISL_17050022, EPI_ISL_17050263, EPI_ISL_17050021, EPI_ISL_17050028,
EPI_ISL_17050027, EPI_ISL_17050269, EPI_ISL_17050268, EPI_ISL_17050026,
EPI_ISL_17050267, EPI_ISL_17050025, EPI_ISL_17050262, EPI_ISL_17050020,
EPI_ISL_17050261, EPI_ISL_17050260, EPI_ISL_17050057, EPI_ISL_17050299,
EPI_ISL_17050056, EPI_ISL_17050298, EPI_ISL_17050055, EPI_ISL_17050297,
EPI_ISL_17050054, EPI_ISL_17050296, EPI_ISL_17050059, EPI_ISL_17050058,
EPI_ISL_17050291, EPI_ISL_17050290, EPI_ISL_17050053, EPI_ISL_17050295,
EPI_ISL_17050052, EPI_ISL_17050294, EPI_ISL_17050051, EPI_ISL_17050293,
EPI_ISL_17050050, EPI_ISL_17050292, EPI_ISL_17050046, EPI_ISL_17050288,

EPI_ISL_17050045, EPI_ISL_17050287, EPI_ISL_17050044, EPI_ISL_17050286,
EPI_ISL_17050043, EPI_ISL_17050285, EPI_ISL_17050049, EPI_ISL_17050048,
EPI_ISL_17050047, EPI_ISL_17050289, EPI_ISL_17050280, EPI_ISL_17050042,
EPI_ISL_17050284, EPI_ISL_17050041, EPI_ISL_17050283, EPI_ISL_17050040,
EPI_ISL_17050282, EPI_ISL_17050281, EPI_ISL_17050079, EPI_ISL_17050078,
EPI_ISL_17050077, EPI_ISL_17050076, EPI_ISL_17050071, EPI_ISL_17050070,
EPI_ISL_17050075, EPI_ISL_17050074, EPI_ISL_17050073, EPI_ISL_17050072,
EPI_ISL_17050068, EPI_ISL_17050067, EPI_ISL_17050066, EPI_ISL_17050065,
EPI_ISL_17050069, EPI_ISL_17050060, EPI_ISL_17050064, EPI_ISL_17050063,
EPI_ISL_17050062, EPI_ISL_17050061, EPI_ISL_17050417, EPI_ISL_17050659,
EPI_ISL_17050416, EPI_ISL_17050658, EPI_ISL_17050415, EPI_ISL_17050657,
EPI_ISL_17050414, EPI_ISL_17050656, EPI_ISL_17050419, EPI_ISL_17050418,
EPI_ISL_17050651, EPI_ISL_17050650, EPI_ISL_17050413, EPI_ISL_17050655,
EPI_ISL_17050412, EPI_ISL_17050654, EPI_ISL_17050411, EPI_ISL_17050653,
EPI_ISL_17050410, EPI_ISL_17050652, EPI_ISL_17050406, EPI_ISL_17050648,
EPI_ISL_17050405, EPI_ISL_17050647, EPI_ISL_17050404, EPI_ISL_17050646,
EPI_ISL_17050403, EPI_ISL_17050645, EPI_ISL_17050409, EPI_ISL_17050408,
EPI_ISL_17050407, EPI_ISL_17050649, EPI_ISL_17050640, EPI_ISL_17050402,
EPI_ISL_17050644, EPI_ISL_17050401, EPI_ISL_17050643, EPI_ISL_17050400,
EPI_ISL_17050642, EPI_ISL_17050641, EPI_ISL_17050439, EPI_ISL_17050438,
EPI_ISL_17050437, EPI_ISL_17050679, EPI_ISL_17050436, EPI_ISL_17050678,
EPI_ISL_17050431, EPI_ISL_17050673, EPI_ISL_17050430, EPI_ISL_17050672,
EPI_ISL_17050671, EPI_ISL_17050670, EPI_ISL_17050435, EPI_ISL_17050677,
EPI_ISL_17050434, EPI_ISL_17050676, EPI_ISL_17050433, EPI_ISL_17050675,
EPI_ISL_17050432, EPI_ISL_17050674, EPI_ISL_17050428, EPI_ISL_17050427,
EPI_ISL_17050669, EPI_ISL_17050426, EPI_ISL_17050668, EPI_ISL_17050425,
EPI_ISL_17050667, EPI_ISL_17050429, EPI_ISL_17050420, EPI_ISL_17050662,
EPI_ISL_17050661, EPI_ISL_17050660, EPI_ISL_17050424, EPI_ISL_17050666,
EPI_ISL_17050423, EPI_ISL_17050665, EPI_ISL_17050422, EPI_ISL_17050664,
EPI_ISL_17050421, EPI_ISL_17050663, EPI_ISL_17050219, EPI_ISL_17050218,
EPI_ISL_17050217, EPI_ISL_17050459, EPI_ISL_17050458, EPI_ISL_17050216,
EPI_ISL_17050453, EPI_ISL_17050211, EPI_ISL_17050695, EPI_ISL_17050452,
EPI_ISL_17050210, EPI_ISL_17050694, EPI_ISL_17050451, EPI_ISL_17050693,
EPI_ISL_17050450, EPI_ISL_17050692, EPI_ISL_17050457, EPI_ISL_17050215,
EPI_ISL_17050699, EPI_ISL_17050456, EPI_ISL_17050214, EPI_ISL_17050698,
EPI_ISL_17050455, EPI_ISL_17050213, EPI_ISL_17050697, EPI_ISL_17050454,
EPI_ISL_17050212, EPI_ISL_17050696, EPI_ISL_17050691, EPI_ISL_17050690,

EPI_ISL_17050208, EPI_ISL_17050207, EPI_ISL_17050449, EPI_ISL_17050448,
EPI_ISL_17050206, EPI_ISL_17050447, EPI_ISL_17050205, EPI_ISL_17050689,
EPI_ISL_17050209, EPI_ISL_17050442, EPI_ISL_17050200, EPI_ISL_17050684,
EPI_ISL_17050441, EPI_ISL_17050683, EPI_ISL_17050440, EPI_ISL_17050682,
EPI_ISL_17050681, EPI_ISL_17050446, EPI_ISL_17050204, EPI_ISL_17050688,
EPI_ISL_17050445, EPI_ISL_17050203, EPI_ISL_17050687, EPI_ISL_17050444,
EPI_ISL_17050202, EPI_ISL_17050686, EPI_ISL_17050443, EPI_ISL_17050201,
EPI_ISL_17050685, EPI_ISL_17050680, EPI_ISL_17050239, EPI_ISL_17050238,
EPI_ISL_17050475, EPI_ISL_17050233, EPI_ISL_17050474, EPI_ISL_17050232,
EPI_ISL_17050473, EPI_ISL_17050231, EPI_ISL_17050472, EPI_ISL_17050230,
EPI_ISL_17050237, EPI_ISL_17050479, EPI_ISL_17050478, EPI_ISL_17050236,
EPI_ISL_17050477, EPI_ISL_17050235, EPI_ISL_17050476, EPI_ISL_17050234,
EPI_ISL_17050471, EPI_ISL_17050470, EPI_ISL_17050229, EPI_ISL_17050228,
EPI_ISL_17050227, EPI_ISL_17050469, EPI_ISL_17050464, EPI_ISL_17050222,
EPI_ISL_17050463, EPI_ISL_17050221, EPI_ISL_17050462, EPI_ISL_17050220,
EPI_ISL_17050461, EPI_ISL_17050468, EPI_ISL_17050226, EPI_ISL_17050467,
EPI_ISL_17050225, EPI_ISL_17050466, EPI_ISL_17050224, EPI_ISL_17050465,
EPI_ISL_17050223, EPI_ISL_17050460, EPI_ISL_17050615, EPI_ISL_17050614,
EPI_ISL_17050613, EPI_ISL_17050612, EPI_ISL_17050619, EPI_ISL_17050618,
EPI_ISL_17050617, EPI_ISL_17050616, EPI_ISL_17050611, EPI_ISL_17050610,
EPI_ISL_17050609, EPI_ISL_17050604, EPI_ISL_17050603, EPI_ISL_17050602,
EPI_ISL_17050601, EPI_ISL_17050608, EPI_ISL_17050607, EPI_ISL_17050606,
EPI_ISL_17050605, EPI_ISL_17050600, EPI_ISL_17050637, EPI_ISL_17050636,
EPI_ISL_17050635, EPI_ISL_17050634, EPI_ISL_17050639, EPI_ISL_17050638,
EPI_ISL_17050633, EPI_ISL_17050632, EPI_ISL_17050631, EPI_ISL_17050630,
EPI_ISL_17050626, EPI_ISL_17050625, EPI_ISL_17050624, EPI_ISL_17050623,
EPI_ISL_17050629, EPI_ISL_17050628, EPI_ISL_17050627, EPI_ISL_17050622,
EPI_ISL_17050621.

Figura 2S. Questionário demográfico e clínico aplicado a todos os pacientes que se cadastraram na Plataforma SC. Dados de contatos próximos, fator de risco, também foram registrados (registro da Plataforma SC).



Admin São Caetano do Sul - Fase 1

Questionário inicial

Nome do Paciente

Celular :

Idade :

Data nascimento :

Sexo:

Rua:

Número:

Complemento:

CEP:

CPF:

Cartão SUS:

CID CARD:

Status pessoa: **contato inicial**

Id do estudo:

IMC:

2 - Dados demográficos

2-1: Se do sexo feminino, está grávida?	<input type="radio"/> Não <input type="radio"/> Sim <input type="radio"/> Não se aplica
---	---

2-2: A partir do dia 1º de Março você atuou em alguma destas áreas:	<input type="radio"/> Não <input type="radio"/> Profissional da saúde <input type="radio"/> Áreas essenciais (segurança, bombeiro, farmácia, supermercado, transporte público) <input type="radio"/> Cuidador (a)
2-3: Nível Educacional:	<input type="radio"/> Nunca fui a escola <input type="radio"/> Ensino Fundamental incompleto <input type="radio"/> Ensino Fundamental (1º a 8º) <input type="radio"/> Ensino Médio (1º a 3º) ou Técnico completo <input type="radio"/> Curso Superior Completo
2-4: Peso (kg):	<input type="text"/>
2-5: Altura (m):	<input type="text"/>

3 - Dados clínicos

3-1: Teve febre?	<input type="radio"/> Não <input type="radio"/> Sim
3-4: Você tem tosse?	<input type="radio"/> Não <input type="radio"/> Sim
3-6: Você tem dor de garganta?	<input type="radio"/> Não <input type="radio"/> Sim
3-7: Você está com o nariz entupido?	<input type="radio"/> Não <input type="radio"/> Sim
3-8: Você está com o nariz escorrendo?	<input type="radio"/> Não <input type="radio"/> Sim
3-9: Você tem dor de cabeça?	<input type="radio"/> Não <input type="radio"/> Sim
3-10: Você tem cansaço?	<input type="radio"/> Não <input type="radio"/> Sim

3-11: Você tem fraqueza?	<input type="radio"/> Não <input type="radio"/> Sim
3-12: Você tem falta de apetite?	<input type="radio"/> Não <input type="radio"/> Sim
3-13: Você tem dor no corpo?	<input type="radio"/> Não <input type="radio"/> Sim
3-15: Você tem dor nas juntas?	<input type="radio"/> Não <input type="radio"/> Sim
3-16: Você tem diarreia?	<input type="radio"/> Não <input type="radio"/> Sim
3-17: Você tem enjoos?	<input type="radio"/> Não <input type="radio"/> Sim
3-18: Você vomitou?	<input type="radio"/> Não <input type="radio"/> Sim
3-19: Você sentiu alteração no olfato (sentir cheiro)?	<input type="radio"/> Não <input type="radio"/> Sim
3-20: Você sentiu alteração no paladar (diferenciar doce de salgado)?	<input type="radio"/> Não <input type="radio"/> Sim
3-21: Há quantos dias você começou a sentir esses sintomas?	<input type="text"/> dias
3-22: Você teve algum contato próximo ou íntimo com alguém nos últimos 5 dias?	<input type="radio"/> Não <input type="radio"/> Sim

4 - Contatos próximos nos últimos 5 dias

Adicionar linha

6 - Fator de risco para evolução desfavorável

6-9: Você teve contacto com uma pessoa que testou positivo para o corona vírus?	<input type="radio"/> Não <input type="radio"/> Sim
6-10: Você é um profissional da saúde?	<input type="radio"/> Não <input type="radio"/> Sim

8 - Orientações

8-4: Foi orientado a buscar o serviço de saúde?	<input type="radio"/> Não <input type="radio"/> Sim
---	---

9. ANEXO I

COMISSÃO NACIONAL DE ÉTICA EM PESQUISA



PARECER CONSUBSTANCIADO DA CONEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: A Diversidade Genética do Coronavírus no Brasil

Pesquisador: Ester Cerdeira Sabino

Área Temática: A critério do CEP

Versão: 2

CAAE: 30101720.1.0000.0068

Instituição Proponente: Faculdade de Medicina da Universidade de São Paulo

Patrocinador Principal: FUNDAÇÃO DE AMPARO A PESQUISA DO ESTADO DE SÃO PAULO

DADOS DO PARECER

Número do Parecer: 3.928.374

Apresentação do Projeto:

As informações contidas nos campos "Apresentação do Projeto", "Objetivo da Pesquisa" e "Avaliação dos Riscos e Benefícios" foram obtidas do documento contendo as Informações Básicas da Pesquisa (PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1525172.pdf de 22/03/2020) e do Projeto Detalhado.

INTRODUÇÃO

Casos de pneumonia de etiologia desconhecida foram descritos em Wuhan, capital da província de Hubei, China desde dezembro de 2019. O patógeno responsável por essa síndrome respiratória foi logo identificado como um novo vírus RNA betacoronavírus relacionado à até então denominada "severe acute respiratory syndrome coronavirus 2" (SARS-CoV-2)¹ sendo a doença decorrente dessa infecção popularmente chamada de Covid-19 (coronavirus disease 2019). Manifestações clínicas encontradas foram predominantemente febre, tosse e cansaço [2] e um número importante de pessoas afetadas necessitou de hospitalização, assistência respiratória e admissão em Unidades de Terapia Intensiva o que causou grande impacto nos serviços de saúde. Embora muitos esforços tenham sido tomados para o controle e a avaliação da dinâmica da epidemia [1,2] ainda existem muitas dúvidas sobre o comportamento da infecção e faltam dados epidemiológicos consistentes o que nos incita a pesquisar como será o comportamento da epidemia de Covid-19 na população brasileira. Podemos descrever temporalmente uma epidemia em diversas fases como

Endereço: SRTVN 701, Via W 5 Norte, lote D - Edifício PO 700, 3º andar
Bairro: Asa Norte **CEP:** 70.719-040
UF: DF **Município:** BRASÍLIA
Telefone: (61)3315-5877 **E-mail:** conep@saude.gov.br

COMISSÃO NACIONAL DE
ÉTICA EM PESQUISA



Continuação do Parecer: 3.928.374

introdução, difusão, amplificação e resolução. Cada momento tem sua particularidade na relação entre o patógeno e o ambiente no qual ele se dissemina e o atual momento de introdução e amplificação do Covid-19 é essencial para realização de estudo que propiciem medidas de controle. Avaliar as particularidades da disseminação e obter taxas de frequência como morbidade (relação do número de doentes sobre a população exposta), mortalidade (relação do número de mortos sobre a população exposta), letalidade (relação do número de mortos sobre o número de doentes), porcentagem de portadores assintomáticos capazes de transmitir a doença, soroprevalência na população saudável ainda não foram totalmente elucidados é o primeiro passo para explicar uma epidemia e melhor propor medidas de ação [3, 4]. No momento atual o mundo se prepara com estudos a serem desenhados que permitirão buscar as respostas mais adequadas. A possibilidade de sequenciamento é uma fonte rica de informação epidemiológica, uma vez que as sequências podem inferir as relações epidemiológicas entre os vírus. A recente disponibilidade da tecnologia de sequenciamento portátil MinION da Oxford Nanopore, cujo barateamento permitiu maior acesso ao uso da tecnologia permitindo um processo mais rápido e hábil (como obtido pelo nosso grupo nos primeiros casos de coronavírus no Brasil). Esta informação é crucial no fornecimento de parâmetros importantes sobre a epidemia incluindo tendências e dinâmicas. A presença de diversos subtipos virais e a possibilidade de recombinações, mutações e deleções do vírus do Covid-19 também favorece mudanças no perfil patogênico que podem impactar a epidemia [5, 6]. A avaliação de perfil genético dos vírus circulantes no Brasil é necessária como medida de vigilância afim de entender o perfil viral predominante e compreender a dinâmica epidêmica possibilitando medidas de controle e de respostas em políticas de saúde pública.

HIPÓTESE

O perfil genético do coronavírus ajudará a compreender a epidemia do país.

METODOLOGIA

Serão sequenciadas aproximadamente 500 amostras de acordo com a necessidade de grupos de pesquisa e laboratórios parceiros, no âmbito do CADDE Centro para Descoberta, Diagnóstico, Genômica e Epidemiologia de Arbovírus, utilizando a tecnologia por nanoporos, já empregada nos projetos em desenvolvimento.

Para tanto, serão utilizados os seguintes critérios:

Serão recebidas amostras de RNA dos grupos participantes. As amostras de RNA serão direcionadas para a síntese de cDNA.

Endereço: SRTVN 701, Via W 5 Norte, lote D - Edifício PO 700, 3º andar
Bairro: Asa Norte **CEP:** 70.719-040
UF: DF **Município:** BRASILIA
Telefone: (61)3315-5877 **E-mail:** conep@saude.gov.br

COMISSÃO NACIONAL DE ÉTICA EM PESQUISA



Continuação do Parecer: 3.928.374

- Síntese de cDNA e PCR Multiplex para amplificação do genoma completo do SARS-CoV-2.

O cDNA será reversamente transcrito a partir do RNA viral usando o kit Protoscript II First Strand Sequencing (NEB) com iniciação através de Random Hexamers (Invitrogen). A PCR multiplex será realizada utilizando a enzima Q5 High Fidelity DNA Polymerase (New England Biolabs) e o esquema de primers para sequenciamento descrito e disponibilizado pela rede ARTIC (<https://artic.network/protocols.html>) [7]. As amostras serão submetidas a 35 ciclos de PCR usando as condições de termociclagem e condições de reação descritas no protocolo. Os produtos de PCR serão purificados por beads magnéticas Ampure XP (Beckman Coulter) e as concentrações serão medidas usando o kit Qubit dsDNA High Sensitivity no fluorímetro Qubit (Invitrogen).

- Preparo da Biblioteca de Sequenciamento

As bibliotecas serão preparadas para sequenciamento na plataforma MinION (Oxford Nanopore Technologies). Este protocolo é compatível com o preparo atual da biblioteca de barcodes 1D Kit (EXP-NBD103 / SQK-LSK108). Será então, realizada usando o Ligation Sequencing 1D (SQKLSK108) e o kit Native Barcoding (EXP-NBD103) de acordo com as instruções do nCoV-2019 sequencing protocol (<https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w>) [8]. O DNA amplificado e os controles negativos apropriados serão sequenciados, utilizando células de fluxo FLO-MIN106 específicas para o dispositivo MinION, no qual a corrida será realizada utilizando a abordagem "read until" com auxílio da ferramenta de código aberto RAMPART versão 10.5 usada para atribuir e mapear leituras em tempo real. As novas sequências nucleotídicas do novo coronavírus (SARS-CoV-2) geradas no MinION serão então montadas utilizando pipelines de montagem de genomas disponível em www.artic.network [7].

Em resumo, os arquivos brutos terão sua chamada de bases através do software Guppy, demultiplexados e aparados com Porechop e mapeados contra a cepa de referência Wuhan-Hu-1 (número de acesso GenBank MN908947). As variantes serão chamadas usando Nanopolish 0.11.3.

A partir das sequências geradas, diversas análises de bioinformática podem ser realizadas para compreender a diversidade genética do novo coronavírus circulante no Brasil, entre elas pode-se citar: 1) identificação de SNPs; 2) associação dos SNPs aos genótipos e regiões geográficas; 3) análise do sinal filogenético; 4) análise filogeográfica; 5) análise da pressão seletiva; 6) mapeamento de epítomos; 7) identificação e caracterização dos cluster de transmissão; 8) análise de eventos de recombinação e 9) análises de modelagem comparativa dos produtos protéicos.

CRITÉRIOS DE INCLUSÃO

Endereço: SRTVN 701, Via W 5 Norte, lote D - Edifício PO 700, 3º andar
Bairro: Asa Norte **CEP:** 70.719-040
UF: DF **Município:** BRASILIA
Telefone: (61)3315-5877 **E-mail:** conep@saude.gov.br

COMISSÃO NACIONAL DE
ÉTICA EM PESQUISA



Continuação do Parecer: 3.028.374

Serão sequenciadas as amostras confirmadas como positivas para o SARS-CoV-2 proveniente de pacientes com diagnóstico de COVID-19 realizado pelos pesquisadores participantes.

CRITÉRIOS DE EXCLUSÃO

Amostras não amplificadas.

Objetivo da Pesquisa:

OBJETIVO PRIMÁRIO

O objetivo geral deste projeto é compreender a diversidade viral durante a epidemia de Coronavírus no Brasil.

OBJETIVO SECUNDÁRIO

Verificar se as mutações presentes nas cepas circulantes no Brasil podem alterar a sensibilidade aos novos antirretrovirais.

Avaliação dos Riscos e Benefícios:

RISCOS

Toda e qualquer informação relacionada aos pacientes de onde provém as amostras será codificada inviabilizando a identificação do paciente e a revelação de seu diagnóstico, diminuindo assim o risco de expor informações pessoais e respeita-se o direito ao sigilo do paciente. As informações provenientes desse estudo serão divulgadas prioritariamente aos órgãos responsáveis e meio científico a fim de gerar informações úteis para ação em saúde pública e não causar nenhum prejuízo ou repercussão negativa.

BENEFÍCIOS

O entendimento do perfil genético do SARS-CoV-2 no Brasil permitirá entender melhor a dinâmica epidemiológica da doença no país, quais suas origens e portas de entrada e as tendências epidêmicas. Isso pode resultar em melhorias nas políticas públicas de contenção de surtos e epidemias e elaboração de fluxos para controle da doença em nível mundial. A possibilidade de utilizar tecnologia mais barata para sequenciamento durante evento epidêmico representa transferência de tecnologia de ponta, formação, capacitação técnica e desenvolvimento do campo de epidemiologia genética de surtos e infecções num país de dimensões continentais e com histórico importante de surtos e epidemias.

Endereço: SRTVN 701, Via W 5 Norte, lote D - Edifício PO 700, 3º andar
Bairro: Asa Norte **CEP:** 70.719-040
UF: DF **Município:** BRASÍLIA
Telefone: (61)3315-5877 **E-mail:** conep@saude.gov.br

10. ANEXO II

5134 - 9888251 / 2 - Flavia Cristina da Silva Sales

Email: flaviasalles@usp.br

Data de Nascimento: 24/03/1989

Cédula de Identidade: RG - 35.074.091-4 - SP

Local de Nascimento: Estado de São Paulo

Nacionalidade: Brasileira

Graduação: Bacharela em Biomedicina - Centro Universitário das Faculdades Metropolitanas Unidas - São Paulo - Brasil - 2014

Curso: Doutorado Direto

Programa: Doenças Infecciosas e Parasitárias

Data de Matrícula: 19/09/2018

Início da Contagem de Prazo: 19/09/2018

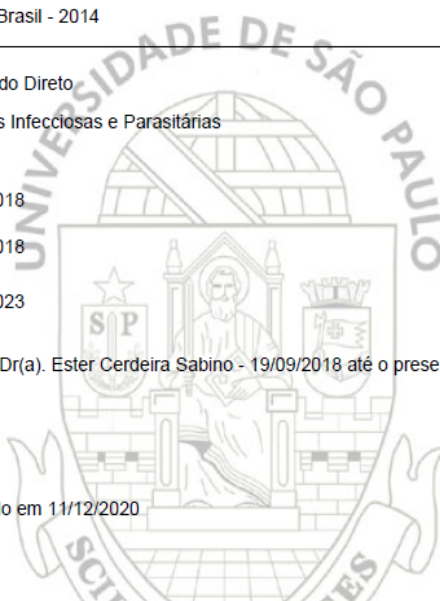
Data Limite para o Depósito: 20/03/2023

Orientador: Prof(a). Dr(a). Ester Cerdeira Sabino - 19/09/2018 até o presente Email: sabinoec@gmail.com

Proficiência em Línguas:

Inglês, Aprovado em 19/09/2018

Data de Aprovação no Exame de Qualificação: Aprovado em 11/12/2020



Sigla	Nome da Disciplina	Início	Término	Carga Horária	Cred.	Freq.	Conc.	Exc.	Situação
PSP5119-1/4	Princípios da Epidemiologia (Faculdade de Saúde Pública - Universidade de São Paulo)	12/03/2019	20/05/2019	90	6	100	B	N	Concluída
IMT5111-2/1	Tópicos avançados em Medicina Tropical (Instituto de Medicina Tropical de São Paulo - Universidade de São Paulo)	22/03/2019	04/07/2019	30	2	100	A	N	Concluída
ICB5747-2/4	Ciências Ômicas em Doenças Infecciosas (CODI) (Instituto de Ciências Biomédicas - Universidade de São Paulo)	01/07/2019	21/07/2019	90	0	-	-	N	Matrícula cancelada
Crédito Externo	Tópicos em Bioinformática IV - Epidemiologia De Surto De Dengue: Capacitação, Pesquisa, Vigilância e Divulgação Científica - Universidade Federal de Minas Gerais (1)	19/08/2019	30/08/2019	-	4	-	T	-	-
IMT5124-2/1	Seminários sobre Vírus Persistentes de Importância em Saúde Pública (Instituto de Medicina Tropical de São Paulo)	05/09/2019	27/11/2019	60	4	100	A	N	Concluída

Sigla	Nome da Disciplina	Início	Término	Carga Horária	Cred.	Freq.	Conc.	Exc.	Situação
Paulo - Universidade de São Paulo)									
ENS5920-2/2	Bioestatística (Escola de Enfermagem - Universidade de São Paulo)	13/02/2020	07/05/2020	60	0	-	-	N	Pré-matricula indeferida
FMM5003-1/1	Princípios de Análise de Dados e de Bioestatística	02/03/2020	15/03/2020	60	0	-	-	N	Matricula cancelada
PSP5128-1/1	Métodos e aplicação de análise de dados não paramétricos (Faculdade de Saúde Pública - Universidade de São Paulo)	14/05/2020	02/07/2020	60	0	-	-	N	Turma cancelada
VPS5747-1/2	Introdução à Bioinformática e Genômica Bacteriana (Faculdade de Medicina Veterinária e Zootecnia - Universidade de São Paulo)	09/06/2020	30/06/2020	45	0	-	-	N	Turma cancelada
FMM5003-1/2	Princípios de Análise de Dados e de Bioestatística	03/08/2020	16/08/2020	60	4	100	A	N	Concluída
IBI5086-1/2	Introdução aos Métodos Estatísticos para Bioinformática (Curso Interunidades: Bioinformática - Universidade de São Paulo)	13/08/2020	26/11/2020	120	8	100	A	N	Concluída
IBI5037-2/2	Algoritmos em Bioinformática (Curso Interunidades: Bioinformática - Universidade de São Paulo)	17/08/2020	11/12/2020	120	0	-	-	N	Matricula cancelada
IBI5035-2/5	Biologia Molecular Computacional (Curso Interunidades: Bioinformática - Universidade de São Paulo)	14/09/2020	27/12/2020	120	0	-	-	N	Matricula cancelada
MFT5725-5/1	Pesquisa Clínica: Delineamento e Condução, Processamento de Sinais Biológicos e Tratamento Estatístico	15/09/2020	23/11/2020	90	0	-	-	N	Matricula cancelada
DPG5005-1/1	A Crise das Pandemias e as Oportunidades para a Construção de um Mundo mais Seguro, Menos Desigual e Sustentável (Pró-Reitoria de Pós-Graduação - Universidade de São Paulo)	30/09/2020	16/12/2020	30	0	-	-	S	Matricula cancelada
FMM5006-1/1	Redação Científica	05/10/2020	18/10/2020	60	0	-	-	N	Matricula cancelada
DPG5006-1/1	A Crise das Pandemias e as Oportunidades para a Construção de um Mundo mais Seguro, Menos	22/03/2021	30/05/2021	30	0	-	-	N	Matricula cancelada

Sigla	Nome da Disciplina	Início	Término	Carga Horária	Cred.	Freq.	Conc.	Exc.	Situação
PSP5122-3/1	Desigual e Sustentável - II (Pró-Reitoria de Pós-Graduação - Universidade de São Paulo) Análise de Dados em Epidemiologia Espacial I (Faculdade de Saúde Pública - Universidade de São Paulo)	12/08/2021	01/12/2021	60	4	100	C	N	Concluída

	Créditos mínimos exigidos		Créditos obtidos
	Para exame de qualificação	Para depósito de tese	
Disciplinas:	12	24	32
Estágios:			
Total:	12	24	32

Créditos Atribuídos à Tese: 176

Observações:

1) Disciplinas(s) cursada(s) na(o) Universidade Federal de Minas Gerais. Atribuição de créditos aprovada pelo(a) Comissão de Pós-Graduação em Sessão de 27/01/2020

Conceito a partir de 02/01/1997:

A - Excelente, com direito a crédito; B - Bom, com direito a crédito; C - Regular, com direito a crédito; R - Reprovado; T - Transferência.

Um(1) crédito equivale a 15 horas de atividade programada.

10.1 Publicações

1. Nicolete, Vanessa C.; Rodrigues, Priscila T.; Fernandes, Anderson R.J.; Corder, Rodrigo M.; Tonini, Juliana; Buss, Lewis F.; **Sales, Flávia C.**; Faria, Nuno R.; Sabino, Ester C.; Castro, Marcia C.; Ferreira, Marcelo U. Epidemiology Of Covid-19 After Emergence Of Sars-Cov-2 Gamma Variant, Brazilian Amazon, 2020-2021. *Emerging Infectious Diseases (Online)*, V. 28, P. 709-712, 2022. Citações:6

2. Claro, Ingra Morales; **Da Silva Sales, Flavia Cristina**; Ramundo, Mariana Severo; Candido, Darlan S.; Silva, Camila A.M.; De Jesus, Jaqueline Goes; Manuli, Erika R.; De Oliveira, Cristina Mendes; Scarpelli, Luciano; Campana, Gustavo; Pybus, Oliver G.; Sabino, Ester Cerdeira; Faria, Nuno Rodrigues; Levi, José Eduardo . Local Transmission Of Sars-Cov-2 Lineage B.1.1.7, Brazil, December 2020. *Emerging Infectious Diseases (Online)*, V. 27, P. 970-972, 2021. Citações:37

3.Faria, Nuno R. Mellan, Thomas A. Whittaker, Charles Claro, Ingra M. Candido, Darlan Da S. Mishra, Swapnil Crispim, Myuki A. E. Sales, **Flavia C. S.** Hawryluk, Iwona Mccrone, John T. Hulswit, Ruben J. G. Franco, Lucas A. M. Ramundo, Mariana S. De Jesus, Jaqueline G. Andrade, Pamela S. Coletti, Thais M. Ferreira, Giulia M. Silva, Camila A. M. Manuli, Erika R. Pereira, Rafael H. M. Peixoto, Pedro S. Kraemer, Moritz U. G. Gaburo, Nelson Camilo, Cecilia Da C. Hoeltgebaum, Henrique , Et Al.; *Genomics And Epidemiology Of The P.1 Sars-Cov-2 Lineage In Manaus, Brazil. Science*, V. 372, P. 815-821, 2021. Citações:633

4.Claro, Ingra M.; Ramundo, Mariana S.; Coletti, Thais M.; Da Silva, Camila A. M.; Valenca, Ian N.Cândido, Darlan S.; **Sales, Flavia C. S.**; Manuli, Erika R.; De Jesus, Jaqueline G.; De Paula, Anderson; Felix, Alvina Clara; Andrade, Pamela Dos Santos; Pinho, Mariana C.; Souza, William M.; Amorim, Mariene R.; Proenca-Modena, José Luiz; Kallas, Esper G.; Levi, José Eduardo; Faria, Nuno Rodrigues; Sabino, Ester C.; Loman, Nicholas J.; Quick, Joshua . Rapid Viral Metagenomics Using Smart-9n Amplification And Nanopore Sequencing. *Wellcome Open Research*, V. 6, P. 241, 2021.

5.Giovanetti, Marta Faria, Nuno Rodrigues Lourenço, José Goes De Jesus, Jaqueline Xavier, Joilson Claro, Ingra Morales Kraemer, Moritz U.G. Fonseca, Vagner Dellicour, Simon Thézé, Julien **Da Silva Salles, Flavia** Gräf, Tiago Silveira, Paola Paz Do Nascimento, Valdinete Alves Costa De Souza, Victor De Melo Iani, Felipe Campos Castilho-Martins, Emerson Augusto Cruz, Laura Nogueira Wallau, Gabriel Fabri, Allison Levy, Flávia Quick, Joshua De Azevedo, Vasco Aguiar, Renato Santana De Oliveira, Tulio , Et Al.; *Genomic And*

Epidemiological Surveillance Of Zika Virus In The Amazon Region. Cell Reports, V. 30, P. 2275-2283.E7, 2020. Citações:19

6.De Oliveira, Léa Campos; Pereira, Natalia Bueno; Moreira, Carlos Henrique Valente; Bierrenbach, Ana Luiza; **Salles, Flavia Cristina**; De Souza-Basqueira, Marcela; Manuli, Erika Regina; Ferreira, Ariela Mota; Oliveira, Cláudia Di Lorenzo; Cardoso, Clareci Silva; Ribeiro, Antonio Luiz P.; Sabino, Ester Cerdeira . Elisa Saliva For Trypanosoma Cruzi Antibody Detection: An Alternative For Serological Surveys In Endemic Regions. American Journal Of Tropical Medicine And Hygiene, V. 102, P. 800-803, 2020. Citações:1

7.De Jesus, Jaqueline Goes; Dutra, Karina Rocha; **Sales, Flavia Cristina Da Silva**; Claro, Ingra Morales; Terzian, Ana Carolina; Candido, Darlan Da Silva; Hill, Sarah C; Thézé, Julien; Torres, Celeste; D'Agostini, Tatiana Lang; Felix, Alvina Clara; Reis, Andreia F Negri; Alcantara, Luiz Carlos Junior; De Abreu, André L; Croda, Júlio Hr; De Oliveira, Wanderson K; De Filipis, Ana Maria Bispo; Camis, Maria Do Carmo Rodrigues Dos Santos; Romano, Camila Malta; Loman, Nick J; Pybus, Oliver G; Sabino, Ester Cerdeira; Nogueira, Mauricio L; Faria, Nuno Rodrigues . Genomic Detection Of A Virus Lineage Replacement Event Of Dengue Virus Serotype 2 In Brazil, 2019. Memórias Do Instituto Oswaldo Cruz, V. 115, P. E190423, 2020. Citações:9

8. Jesus, Jaqueline Goes De; Sacchi, Claudio; Candido, Darlan Da Silva; Claro, Ingra Morales; Sales, **Flávia Cristina Silva**; Manuli, Erika Regina; Silva, Daniela Bernardes Borges Da; Paiva, Terezinha Maria De; Pinho, Margarete Aparecida Benega; Santos, Katia Correa De Oliveira; Hill, Sarah Catherine; Aguiar, Renato Santana; Romero, Filipe; Santos, Fabiana Cristina Pereira Dos; Gonçalves, Claudia Regina; Timenetsky, Maria Do Carmo; Quick, Joshua; Croda, Julio Henrique Rosa; Oliveira, Wanderson De; Rambaut, Andrew; Pybus, Oliver G.; Loman, Nicholas J.; Sabino, Ester Cerdeira; Faria, Nuno Rodrigues . Importation And Early Local Transmission Of Covid-19 In Brazil, 2020. Revista Do Instituto De Medicina Tropical De São Paulo, V. 62, P. 62:E30, 2020. Citações:35

9.Candido, Darlan S. Claro, Ingra M. De Jesus, Jaqueline G. Souza, William M. Moreira, Filipe R. R. Dellicour, Simon Mellan, Thomas A. Du Plessis, Louis Pereira, Rafael H. M. **Sales, Flavia C. S.** Manuli, Erika R. Thézé, Julien Almeida, Luiz Menezes, Mariane T. Voloch, Carolina M. Fumagalli, Marcilio J. Coletti, Thaís M. Da Silva, Camila A. M. Ramundo, Mariana S. Amorim, Mariene R. Hoeltgebaum, Henrique H. Mishra, Swapnil Gill, Mandev S. Carvalho, Luiz M. Buss, Lewis F. , Et Al.; Evolution And Epidemic Spread Of Sars-Cov-2 In Brazil. Science, V. 1, P. Eabd2161, 2020. Citações:276

10.De Souza, William Marciel Buss, Lewis Fletcher Candido, Darlan Da Silva

Carrera, Jean-Paul Li, Sabrina Zarebski, Alexander E. Pereira, Rafael Henrique Moraes Prete, Carlos A. De Souza-Santos, Andreza Aruska Parag, Kris V. Belotti, Maria Carolina T. D. Vincenti-Gonzalez, Maria F. Messina, Janey Da Silva Sales, Flavia Cristina Andrade, Pamela Dos Santos Nascimento, Vítor Heloiz Ghilardi, Fabio Abade, Leandro Gutierrez, Bernardo Kraemer, Moritz U. G. Braga, Carlos K. V. Aguiar, Renato Santana Alexander, Neal Mayaud, Philippe Brady, Oliver J. , Et Al.; Epidemiological And Clinical Characteristics Of The Covid-19 Epidemic In Brazil. *Nature Human Behaviour*, V. 4, P. 856-865, 2020.Citações:147

11.Hill, Sarah C. De Souza, Renato Thézé, Julien Claro, Ingra Aguiar, Renato S. Abade, Leandro Santos, Fabiana C. P. Cunha, Mariana S. Nogueira, Juliana **S. Salles, Flavia C.** S. Rocco, Iray M. Maeda, Adriana Y. Vasami, Fernanda G. S. Du Plessis, Louis Silveira, Paola P. De Jesus, Jaqueline G. Quick, Joshua Fernandes, Natália C. C. A. Guerra, Juliana M. Réssio, Rodrigo A. Giovanetti, Marta Alcantara, Luiz C. J. Cirqueira, Cinthya S. Díaz-Delgado, Josué Macedo, Fernando L. L. , Et Al.; Genomic Surveillance Of Yellow Fever Virus Epizootic In São Paulo, Brazil, 2016 - 2018. *Plos Pathogens*, V. 16, P. E1008699, 2020. Citações:15

12.Prates, Gabriela; Assone, Tatiane; Corral, Marcelo; Baldassin, Maíra P. M; Mitiko, Tatiane; **Silva Sales, Flávia C.**; Haziot, Michel E.; Smid, Jerusa; Fonseca, Luiz A. M.; Toledo, Fernanda; Penalva De Oliveira, Augusto C.; Casseb, Jorge . Prognosis Markers For Monitoring Htlv-1 Neurologic Disease. *Neurology: Clinical Practice*, V. 1, P. 10.1212/Cpj.00, 2020.Citações:2

13.Buss, Lewis F.; Campos De Oliveira- Da Silva, Léa; Moreira, Carlos H. V.; Manuli, Erika R.; **Sales, Flavia C.**; Morales, Ingra; Di Germanio, Clara; De Almeida-Neto, Cesar; Bakkour, Sonia; Constable, Paul; Pinto-Filho, Marcelo M.; Ribeiro, Antonio L.; Busch, Michael; Sabino, Ester C. . Declining Antibody Levels To Trypanosoma Cruzi Correlate With Polymerase Chain Reaction Positivity And Electrocardiographic Changes In A Retrospective Cohort Of Untreated Brazilian Blood Donors. *Plos Neglected Tropical Diseases*, V. 14, P. E0008787, 2020.Citações:7

14. Naveca, Felipe Gomes Claro, Ingra Giovanetti, Marta De Jesus, Jaqueline Goes Xavier, Joilson Iani, Felipe Campos De Melo Do Nascimento, Valdinete Alves De Souza, Victor Costa Silveira, Paola Paz Lourenço, José Santillana, Mauricio Kraemer, Moritz U. G. Quick, Josh Hill, Sarah C. Thézé, Julien Carvalho, Rodrigo Dias De Oliveira Azevedo, Vasco Salles, **Flavia Cristina Da Silva** Nunes, Márcio Roberto Teixeira Lemos, Poliana Da Silva Candido, Darlan Da Silva Pereira, Glauco De Carvalho Oliveira, Marluce Aparecida Assunção Meneses, Cátia Alexandra Ribeiro Maito, Rodrigo Melo , Et Al.; Genomic, Epidemiological And Digital Surveillance Of Chikungunya Virus In The Brazilian

Amazon. Plos Neglected Tropical Diseases, V. 13, P. E0007065, 2019.
Citações:27

15.Ribeiro, Roberto Marques; Souza-Basqueira, Marcela De; Oliveira, Léa Campos De; **Salles, Flavia Cristina**; Pereira, Natalia Bueno; Sabino, Ester Cerdeira . An Alternative Storage Method For Characterization Of The Intestinal Microbiota Through Next Generation Sequencing. Revista Do Instituto De Medicina Tropical De São Paulo, V. 60, P. On-Line Version, 2018.

Citações:4

11. REFERÊNCIAS BIBLIOGRÁFICAS

1. 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. *Nat Microbiol.* 2020 Apr;5(4):536–44.
2. Chen Y, Liu Q, Guo D. Emerging coronaviruses: genome structure, replication, and pathogenesis. *J Med Virol.* 2020 Feb 7;92(4):418–23.
3. Payne S. Family Coronaviridae. *Viruses.* Elsevier; 2017. p. 149–58.
4. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol.* 2020 Oct 28;19(3):155–70.
5. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol.* 2015;1282:1–23.
6. Peacock TP, Penrice-Randal R, Hiscox JA, Barclay WS. SARS-CoV-2 one year on: evidence for ongoing viral adaptation. *J Gen Virol.* 2021 Apr 15;102(4).
7. Perlman S, Netland J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat Rev Microbiol.* 2009 Jun;7(6):439–50.
8. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science.* 2003 Jun 20;300(5627):1961–6.
9. Hemida MG, Perera RA, Wang P, Alhammadi MA, Siu LY, Li M, et al. Middle East Respiratory Syndrome (MERS) coronavirus seroprevalence in domestic livestock in Saudi Arabia, 2010 to 2013. *Euro Surveill.* 2013 Dec 12;18(50):20659.
10. Müller MA, Corman VM, Jores J, Meyer B, Younan M, Liljander A, et al. MERS coronavirus neutralizing antibodies in camels, Eastern Africa, 1983-1997. *Emerging Infect Dis.* 2014 Dec;20(12):2093–5.
11. Vijgen L, Keyaerts E, Moës E, Thoelen I, Wollants E, Lemey P, et al. Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J Virol.* 2005 Feb;79(3):1595–604.
12. Woo PCY, Lau SKP, Lam CSF, Lau CCY, Tsang AKL, Lau JHN, et al. Discovery of seven novel mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. *J Virol.* 2012

- Apr;86(7):3995–4008.
13. Forni D, Cagliani R, Clerici M, Sironi M. Molecular evolution of human coronavirus genomes. *Trends Microbiol.* 2017 Jan;25(1):35–48.
 14. Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* 2019 Mar;17(3):181–92.
 15. Tong S, Conrardy C, Ruone S, Kuzmin IV, Guo X, Tao Y, et al. Detection of novel SARS-like and other coronaviruses in bats from Kenya. *Emerging Infect Dis.* 2009 Mar;15(3):482–5.
 16. Gloza-Rausch F, Ipsen A, Seebens A, Göttsche M, Panning M, Drexler JF, et al. Detection and prevalence patterns of group I coronaviruses in bats, northern Germany. *Emerging Infect Dis.* 2008 Apr;14(4):626–31.
 17. Dominguez SR, O’Shea TJ, Oko LM, Holmes KV. Detection of group 1 coronaviruses in bats in North America. *Emerging Infect Dis.* 2007 Sep;13(9):1295–300.
 18. Woo PCY, Wang M, Lau SKP, Xu H, Poon RWS, Guo R, et al. Comparative analysis of twelve genomes of three novel group 2c and group 2d coronaviruses reveals unique group and subgroup features. *J Virol.* 2007 Feb;81(4):1574–85.
 19. Woo PCY, Lau SKP, Lam CSF, Lai KKY, Huang Y, Lee P, et al. Comparative analysis of complete genome sequences of three avian coronaviruses reveals a novel group 3c coronavirus. *J Virol.* 2009 Jan;83(2):908–17.
 20. Alkhovsky S, Lenshin S, Romashin A, Vishnevskaya T, Vyshemirsky O, Bulycheva Y, et al. SARS-like Coronaviruses in Horseshoe Bats (*Rhinolophus* spp.) in Russia, 2020. *Viruses.* 2022 Jan 9;14(1).
 21. Dong BQ, Liu W, Fan XH, Vijaykrishna D, Tang XC, Gao F, et al. Detection of a novel and highly divergent coronavirus from asian leopard cats and Chinese ferret badgers in Southern China. *J Virol.* 2007 Jul;81(13):6920–6.
 22. Mihindukulasuriya KA, Wu G, St Leger J, Nordhausen RW, Wang D. Identification of a novel coronavirus from a beluga whale by using a panviral microarray. *J Virol.* 2008 May;82(10):5084–8.
 23. Garbino J, Crespo S, Aubert JD, Rochat T, Ninet B, Deffernez C, et al. A prospective hospital-based study of the clinical impact of non-severe acute respiratory syndrome (Non-SARS)-related human coronavirus infection. *Clin Infect Dis.* 2006 Oct 15;43(8):1009–15.
 24. Fouchier RAM, Hartwig NG, Bestebroer TM, Niemeyer B, de Jong JC, Simon JH, et al. A previously undescribed coronavirus associated with respiratory disease in humans. *Proc Natl Acad Sci USA.* 2004 Apr 20;101(16):6212–6.

25. van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJM, Wolthers KC, et al. Identification of a new human coronavirus. *Nat Med*. 2004 Apr;10(4):368–73.
26. Woo PCY, Lau SKP, Chu C, Chan K, Tsoi H, Huang Y, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J Virol*. 2005 Jan;79(2):884–95.
27. Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol*. 2016 Jun;24(6):490–502.
28. Pyrc K, Dijkman R, Deng L, Jebbink MF, Ross HA, Berkhout B, et al. Mosaic structure of human coronavirus NL63, one thousand years of evolution. *J Mol Biol*. 2006 Dec 15;364(5):964–73.
29. Coronaviruses [Internet]. [cited 2022 Sep 9]. Disponível em: <https://www.microbiologybook.org/virol/corona-port.htm>
30. Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, et al. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science*. 2003 Oct 10;302(5643):276–8.
31. Law HKW, Cheung CY, Ng HY, Sia SF, Chan YO, Luk W, et al. Chemokine up-regulation in SARS-coronavirus-infected, monocyte-derived human dendritic cells. *Blood*. 2005 Oct 1;106(7):2366–74.
32. Peiris JSM, Yuen KY, Osterhaus ADME, Stöhr K. The severe acute respiratory syndrome. *N Engl J Med*. 2003 Dec 18;349(25):2431–41.
33. Lee N, Hui D, Wu A, Chan P, Cameron P, Joynt GM, et al. A major outbreak of severe acute respiratory syndrome in Hong Kong. *N Engl J Med*. 2003 May 15;348(20):1986–94.
34. Scales DC, Green K, Chan AK, Poutanen SM, Foster D, Nowak K, et al. Illness in intensive care staff after brief exposure to severe acute respiratory syndrome. *Emerging Infect Dis*. 2003 Oct;9(10):1205–10.
35. Wong G, Liu W, Liu Y, Zhou B, Bi Y, Gao GF. MERS, SARS, and Ebola: The Role of Super-Spreaders in Infectious Disease. *Cell Host Microbe*. 2015 Oct 14;18(4):398–401.
36. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med*. 2012 Nov 8;367(19):1814–20.
37. World Health Organization. Regional Office for the Eastern Mediterranean. MERS situation update: August 2021. 2021 Aug;
38. WHO. WHO EMRO | MERS situation update, September 2019 |

- MERS-CoV | Epidemic and pandemic diseases [Internet]. [cited 2022 Aug 31]. Disponible em: <http://www.emro.who.int/fr/pandemic-epidemic-diseases/mers-cov/mers-situation-update-september-2019.html>
39. Donnelly CA, Malik MR, Elkholy A, Cauchemez S, Van Kerkhove MD. Worldwide Reduction in MERS Cases and Deaths since 2016. *Emerging Infect Dis.* 2019 Sep 17;25(9):1758–60.
 40. Pneumonia of unknown cause – China [Internet]. [cited 2022 Aug 31]. Disponible em: <https://www.who.int/emergencies/disease-outbreak-news/item/2020-DO-N229>
 41. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 2020 Mar;579(7798):270–3.
 42. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. *Nature.* 2020 Mar;579(7798):265–9.
 43. Novel 2019 coronavirus genome - SARS-CoV-2 coronavirus - Virological [Internet]. [cited 2022 Aug 31]. Disponible em: <https://virological.org/t/novel-2019-coronavirus-genome/319>
 44. WHO statement on novel coronavirus in Thailand [Internet]. [cited 2022 Aug 31]. Disponible em: <https://www.who.int/news/item/13-01-2020-who-statement-on-novel-coronavirus-in-thailand>
 45. Novel Coronavirus – Japan [Internet]. [cited 2022 Sep 1]. Disponible em: <https://www.who.int/emergencies/disease-outbreak-news/item/2020-DO-N236>
 46. First Travel-related Case of 2019 Novel Coronavirus Detected in United States | CDC Online Newsroom | CDC [Internet]. [cited 2022 Sep 1]. Disponible em: <https://www.cdc.gov/media/releases/2020/p0121-novel-coronavirus-travel-case.html>
 47. Bernard Stoecklin S, Rolland P, Silue Y, Mailles A, Campese C, Simondon A, et al. First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020. *Euro Surveill.* 2020 Feb;25(6).
 48. Tian H, Liu Y, Li Y, Wu C-H, Chen B, Kraemer MUG, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science.* 2020 May 8;368(6491):638–42.

49. Archived: WHO Timeline - COVID-19 [Internet]. [cited 2022 Aug 31]. Disponible em: <https://www.who.int/news/item/27-04-2020-who-timeline---covid-19>
50. Timeline of WHO's response to COVID-19 [Internet]. [cited 2022 Sep 1]. Disponible em: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline>
51. Lu J, du Plessis L, Liu Z, Hill V, Kang M, Lin H, et al. Genomic Epidemiology of SARS-CoV-2 in Guangdong Province, China. *Cell*. 2020 May 28;181(5):997-1003.e9.
52. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet*. 2008 Apr;9(4):267–76.
53. Sanjuán R, Domingo-Calap P. Mechanisms of viral mutation. *Cell Mol Life Sci*. 2016 Dec;73(23):4433–48.
54. Li G, Fan Y, Lai Y, Han T, Li Z, Zhou P, et al. Coronavirus infections and immune responses. *J Med Virol*. 2020 Apr;92(4):424–32.
55. Long B, Carius BM, Chavez S, Liang SY, Brady WJ, Koyfman A, et al. Clinical update on COVID-19 for the emergency clinician: Presentation and evaluation. *Am J Emerg Med*. 2022 Apr;54:46–57.
56. Guerra FM, Bolotin S, Lim G, Heffernan J, Deeks SL, Li Y, et al. The basic reproduction number (R0) of measles: a systematic review. *Lancet Infect Dis*. 2017 Dec;17(12):e420–8.
57. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020 Feb 15;395(10223):497–506.
58. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020 Mar 28;395(10229):1054–62.
59. Brandal LT, MacDonald E, Veneti L, Ravlo T, Lange H, Naseer U, et al. Outbreak caused by the SARS-CoV-2 Omicron variant in Norway, November to December 2021. *Euro Surveill*. 2021 Dec;26(50).
60. de Souza WM, Buss LF, Candido D da S, Carrera J-P, Li S, Zarebski AE, et al. Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. *Nat Hum Behav*. 2020 Aug;4(8):856–65.
61. Zhang B, Zhou X, Qiu Y, Song Y, Feng F, Feng J, et al. Clinical characteristics of 82 cases of death from COVID-19. *PLoS ONE*. 2020 Jul 9;15(7):e0235458.
62. Zheng X, Wang H, Su Z, Li W, Yang D, Deng F, et al. Co-infection of

- SARS-CoV-2 and Influenza virus in Early Stage of the COVID-19 Epidemic in Wuhan, China. *J Infect*. 2020 Aug;81(2):e128–9.
63. Cuadrado-Payán E, Montagud-Marrahi E, Torres-Elorza M, Bodro M, Blasco M, Poch E, et al. SARS-CoV-2 and influenza virus co-infection. *Lancet*. 2020 May 16;395(10236):e84.
 64. Rawson TM, Moore LSP, Zhu N, Ranganathan N, Skolimowska K, Gilchrist M, et al. Bacterial and Fungal Coinfection in Individuals With Coronavirus: A Rapid Review To Support COVID-19 Antimicrobial Prescribing. *Clin Infect Dis*. 2020 Dec 3;71(9):2459–68.
 65. National SARS-CoV-2 Serology Assay Evaluation Group. Performance characteristics of five immunoassays for SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Infect Dis*. 2020 Dec;20(12):1390–400.
 66. Hanson KE, Caliendo AM, Arias CA, Hayden MK, Englund JA, Lee MJ, et al. The Infectious Diseases Society of America Guidelines on the Diagnosis of COVID-19: Molecular Diagnostic Testing. *Clin Infect Dis*. 2021 Jan 22;
 67. Wang R, Hozumi Y, Yin C, Wei G-W. Decoding SARS-CoV-2 Transmission and Evolution and Ramifications for COVID-19 Diagnosis, Vaccine, and Medicine. *J Chem Inf Model*. 2020 Dec 28;60(12):5853–65.
 68. Hanson KE, Altayar O, Caliendo AM, Arias CA, Englund JA, Hayden MK, et al. The Infectious Diseases Society of America Guidelines on the Diagnosis of COVID-19: Antigen Testing. *Clin Infect Dis*. 2021 Jun 23;
 69. Brümmer LE, Katzenschlager S, Gaeddert M, Erdmann C, Schmitz S, Bota M, et al. Accuracy of novel antigen rapid diagnostics for SARS-CoV-2: A living systematic review and meta-analysis. *PLoS Med*. 2021 Aug 12;18(8):e1003735.
 70. Smith EC, Sexton NR, Denison MR. Thinking outside the triangle: replication fidelity of the largest RNA viruses. *Annu Rev Virol*. 2014 Nov;1(1):111–32.
 71. Tao K, Tzou PL, Nouhin J, Gupta RK, de Oliveira T, Kosakovsky Pond SL, et al. The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat Rev Genet*. 2021 Dec;22(12):757–73.
 72. Jackson B, Boni MF, Bull MJ, Collieran A, Colquhoun RM, Darby AC, et al. Generation and transmission of interlineage recombinants in the SARS-CoV-2 pandemic. *Cell*. 2021 Sep 30;184(20):5179-5188.e8.
 73. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*. 2018 Dec 1;34(23):4121–3.
 74. Virological - Discussion forum for analysis of virus genomes [Internet]. [cited 2023 Mar 14]. Disponível em: <https://virological.org/>

75. Morel B, Barbera P, Czech L, Bettisworth B, Hübner L, Lutteropp S, et al. Phylogenetic Analysis of SARS-CoV-2 Data Is Difficult. *Mol Biol Evol.* 2021 May 4;38(5):1777–91.
76. Maponga TG, Jeffries M, Tegally H, Sutherland AD, Wilkinson E, Lessells R, et al. Persistent SARS-CoV-2 Infection with Accumulation of Mutations in a Patient with Poorly Controlled HIV Infection. *SSRN Journal.* 2022;
77. Koelle K, Martin MA, Antia R, Lopman B, Dean NE. The changing epidemiology of SARS-CoV-2. *Science.* 2022 Mar 11;375(6585):1116–21.
78. Rambaut A, Holmes EC, O’Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol.* 2020 Nov;5(11):1403–7.
79. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell.* 2020 Aug 20;182(4):812-827.e19.
80. WHO WHO. Tracking SARS-CoV-2 variants [Internet]. [cited 2022 Sep 27]. Disponível em: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>
81. Aksamentov I, Roemer C, Hodcroft E, Neher R. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *JOSS.* 2021 Nov 30;6(67):3773.
82. Bal A, Destras G, Gaymard A, Stefic K, Marlet J, Eymieux S, et al. Two-step strategy for the identification of SARS-CoV-2 variant of concern 202012/01 and other variants with spike deletion H69-V70, France, August to December 2020. *Euro Surveill.* 2021 Jan;26(3).
83. Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dingens AS, et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell.* 2020 Sep 3;182(5):1295-1310.e20.
84. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science.* 2021 Apr 9;372(6538).
85. Volz E, Mishra S, Chand M, Barrett JC, The COVID-19 Genomics UK (COG-UK) consortium, Johnson R, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature.* 2021 Mar 25;
86. Claro IM, da Silva Sales FC, Ramundo MS, Candido DS, Silva CAM, de Jesus JG, et al. Local Transmission of SARS-CoV-2 Lineage B.1.1.7, Brazil, December 2020. *Emerging Infect Dis.* 2021 Mar;27(3):970–2.

87. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature*. 2021;592(7854):438–43.
88. Abdool Karim SS, de Oliveira T. New SARS-CoV-2 Variants - Clinical, Public Health, and Vaccine Implications. *N Engl J Med*. 2021 May 13;384(19):1866–8.
89. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings - SARS-CoV-2 coronavirus / nCoV-2019 Genomic Epidemiology - Virological [Internet]. [cited 2022 Sep 30]. Disponível em: <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-co-v-2-lineage-in-manaus-preliminary-findings/586>
90. Messages K. Brief report: New Variant Strain of SARS-CoV-2 Identified in Travelers from Brazil.
91. Candido DS, Claro IM, de Jesus JG, Souza WM, Moreira FRR, Dellicour S, et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science*. 2020 Sep 4;369(6508):1255–60.
92. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido D da S, Mishra S, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*. 2021 May 21;372(6544):815–21.
93. Sabino EC, Buss LF, Carvalho MPS, Prete CA, Crispim MAE, Fraiji NA, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet*. 2021 Feb 6;397(10273):452–5.
94. Dhar MS, Marwal R, Vs R, Ponnusamy K, Jolly B, Bhoyar RC, et al. Genomic characterization and epidemiology of an emerging SARS-CoV-2 variant in Delhi, India. *Science*. 2021 Nov 19;374(6570):995–9.
95. Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, et al. SARS-CoV-2 Spike Mutations, L452R, T478K, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India. *Microorganisms*. 2021 Jul 20;9(7).
96. Mlcochova P, Kemp SA, Dhar MS, Papa G, Meng B, Ferreira IATM, et al. SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature*. 2021 Nov;599(7883):114–9.
97. outbreak.info [Internet]. [cited 2022 Oct 1]. Disponível em: <https://outbreak.info/>
98. Scott L, Hsiao N-Y, Moyo S, Singh L, Tegally H, Dor G, et al. Track Omicron's spread with molecular data. *Science*. 2021 Dec 17;374(6574):1454–5.
99. Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et

- al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature*. 2022 Mar;603(7902):679–86.
100. Liu L, Iketani S, Guo Y, Chan JF-W, Wang M, Liu L, et al. Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. *Nature*. 2022 Feb;602(7898):676–81.
 101. Chemaitelly H, Tang P, Hasan MR, AlMukdad S, Yassine HM, Benslimane FM, et al. Waning of BNT162b2 Vaccine Protection against SARS-CoV-2 Infection in Qatar. *N Engl J Med*. 2021 Dec 9;385(24):e83.
 102. Bar-On YM, Goldberg Y, Mandel M, Bodenheimer O, Freedman L, Kalkstein N, et al. Protection of BNT162b2 Vaccine Booster against Covid-19 in Israel. *N Engl J Med*. 2021 Oct 7;385(15):1393–400.
 103. Jesus JG de, Sacchi C, Candido D da S, Claro IM, Sales FCS, Manuli ER, et al. Importation and early local transmission of COVID-19 in Brazil, 2020. *Rev Inst Med Trop Sao Paulo*. 2020 May 11;62:e30.
 104. Coronavirus Disease (COVID-19) Situation Reports [Internet]. [cited 2022 Sep 1]. Disponível em: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>
 105. www.who.int/docs/default-source/coronaviruse/situation-reports/20200304-sitrep-44-covid-19.pdf.
 106. Candido DDS, Watts A, Abade L, Kraemer MUG, Pybus OG, Croda J, et al. Routes for COVID-19 importation in Brazil. *J Travel Med*. 2020 May 18;27(3).
 107. Faria de Moura Villela E, López RVM, Sato APS, de Oliveira FM, Waldman EA, Van den Bergh R, et al. COVID-19 outbreak in Brazil: adherence to national preventive measures and impact on people's lives, an online survey. *BMC Public Health*. 2021 Jan 18;21(1):152.
 108. The Lancet. COVID-19 in Latin America: a humanitarian crisis. *Lancet*. 2020 Nov 7;396(10261):1463.
 109. Burki T. COVID-19 in Latin America. *Lancet Infect Dis*. 2020 May;20(5):547–8.
 110. www.who.int/docs/default-source/coronaviruse/situation-reports/20200522-covid-19-sitrep-123.pdf.
 111. Whittaker C, Walker PGT, Alhaffar M, Hamlet A, Djaafara BA, Ghani A, et al. Under-reporting of deaths limits our understanding of true burden of covid-19. *BMJ*. 2021 Oct 12;375:n2239.
 112. Naveca FG, Nascimento V, de Souza VC, Corado A de L, Nascimento F,

- Silva G, et al. COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence. *Nat Med*. 2021 Jul;27(7):1230–8.
113. Sant'Anna FH, Muterle Varela AP, Prichula J, Comerlato J, Comerlato CB, Roglio VS, et al. Emergence of the novel SARS-CoV-2 lineage VUI-NP13L and massive spread of P.2 in South Brazil. *Emerg Microbes Infect*. 2021 Dec;10(1):1431–40.
 114. Giovanetti M, Fonseca V, Wilkinson E, Tegally H, San EJ, Althaus CL, et al. Replacement of the Gamma by the Delta variant in Brazil: Impact of lineage displacement on the ongoing pandemic. *Virus Evol*. 2022 Mar 18;8(1):veac024.
 115. Gangavarapu K, Latif AA, Mullen JL, Alkuzweny M, Hufbauer E, Tsueng G, et al. Outbreak.info genomic reports: scalable and dynamic surveillance of SARS-CoV-2 variants and mutations. *Nat Methods*. 2023 Feb 23;
 116. Ferguson N. Report 49: Growth and immune escape of the Omicron SARS-CoV-2 variant of concern in England. Imperial College London. 2021;
 117. Louten J. Virus transmission and epidemiology. *Essential Human Virology*. Elsevier; 2016. p. 71–92.
 118. Kraemer MUG, Pybus OG, Fraser C, Cauchemez S, Rambaut A, Cowling BJ. Monitoring key epidemiological parameters of SARS-CoV-2 transmission. *Nat Med*. 2021 Nov;27(11):1854–5.
 119. Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the basic reproduction number (R_0). *Emerging Infect Dis*. 2019 Jan;25(1):1–4.
 120. Buss LF, Prete CA, Abraham CMM, Mendrone A, Salomon T, de Almeida-Neto C, et al. Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. *Science*. 2021 Jan 15;371(6526):288–92.
 121. Organization WH. Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health, 8 January 2021. World Health Organization; 2021.
 122. Gardy JL, Loman NJ. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat Rev Genet*. 2018 Jan;19(1):9–20.
 123. Khare S, Gurry C, Freitas L, Schultz MB, Bach G, Diallo A, et al. Gisaid's role in pandemic response. *China CDC Wkly*. 2021 Dec 3;3(49):1049–51.
 124. Mena I, Nelson MI, Quezada-Monroy F, Dutta J, Cortes-Fernández R, Lara-Puente JH, et al. Origins of the 2009 H1N1 influenza pandemic in

- swine in Mexico. *eLife*. 2016 Jun 28;5.
125. Ladner JT, Wiley MR, Mate S, Dudas G, Prieto K, Lovett S, et al. Evolution and Spread of Ebola Virus in Liberia, 2014-2015. *Cell Host Microbe*. 2015 Dec 9;18(6):659–69.
 126. Tong Y-G, Shi W-F, Liu D, Qian J, Liang L, Bo X-C, et al. Genetic diversity and evolutionary dynamics of Ebola virus in Sierra Leone. *Nature*. 2015 Aug 6;524(7563):93–6.
 127. Holmes EC, Dudas G, Rambaut A, Andersen KG. The evolution of Ebola virus: Insights from the 2013-2016 epidemic. *Nature*. 2016 Oct 13;538(7624):193–200.
 128. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature*. 2016 Feb 11;530(7589):228–32.
 129. Faria NR, Quick J, Claro IM, Thézé J, de Jesus JG, Giovanetti M, et al. Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature*. 2017 Jun 15;546(7658):406–10.
 130. Xavier J, Giovanetti M, Fonseca V, Thézé J, Gräf T, Fabri A, et al. Circulation of chikungunya virus East/Central/South African lineage in Rio de Janeiro, Brazil. *PLoS ONE*. 2019 Jun 11;14(6):e0217871.
 131. Naveca FG, Claro I, Giovanetti M, de Jesus JG, Xavier J, Iani FC de M, et al. Genomic, epidemiological and digital surveillance of Chikungunya virus in the Brazilian Amazon. *PLoS Negl Trop Dis*. 2019 Mar 7;13(3):e0007065.
 132. de Lima STS, de Souza WM, Cavalcante JW, da Silva Candido D, Fumagalli MJ, Carrera J-P, et al. Fatal outcome of chikungunya virus infection in Brazil. *Clin Infect Dis*. 2021 Oct 5;73(7):e2436–43.
 133. de Jesus JG, Dutra KR, Sales FC da S, Claro IM, Terzian AC, Candido D da S, et al. Genomic detection of a virus lineage replacement event of dengue virus serotype 2 in Brazil, 2019. *Mem Inst Oswaldo Cruz*. 2020 May 15;115:e190423.
 134. Hill SC, de Souza R, Thézé J, Claro I, Aguiar RS, Abade L, et al. Genomic Surveillance of Yellow Fever Virus Epizootic in São Paulo, Brazil, 2016 - 2018. *PLoS Pathog*. 2020 Aug 7;16(8):e1008699.
 135. Grubaugh ND, Ladner JT, Lemey P, Pybus OG, Rambaut A, Holmes EC, et al. Tracking virus outbreaks in the twenty-first century. *Nat Microbiol*. 2019 Jan;4(1):10–9.
 136. Hall MD, Woolhouse MEJ, Rambaut A. Using genomics data to reconstruct transmission trees during disease outbreaks. *Rev - Off Int Epizoot*. 2016 Apr;35(1):287–96.

137. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 2012 Jul 30;9(8):772.
138. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017 Jun;14(6):587–9.
139. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 2012 Aug;29(8):1969–73.
140. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. *PLoS Biol*. 2006 May;4(5):e88.
141. Dudas G, Carvalho LM, Rambaut A, Bedford T. MERS-CoV spillover at the camel-human interface. *eLife*. 2018 Jan 16;7.
142. Faria NR, Kraemer MUG, Hill SC, Goes de Jesus J, Aguiar RS, Iani FCM, et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. *Science*. 2018 Aug 31;361(6405):894–9.
143. Hedge J, Lycett SJ, Rambaut A. Real-time characterization of the molecular epidemiology of an influenza pandemic. *Biol Lett*. 2013 Oct 23;9(5):20130331.
144. Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*. 2004 Jan 16;303(5656):327–32.
145. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog*. 2014 Feb 20;10(2):e1003932.
146. Holmes EC. The phylogeography of human viruses. *Mol Ecol*. 2004 Apr;13(4):745–56.
147. Holmes EC. Evolutionary history and phylogeography of human viruses. *Annu Rev Microbiol*. 2008;62:307–28.
148. Baele G, Suchard MA, Rambaut A, Lemey P. Emerging concepts of data integration in pathogen phylodynamics. *Syst Biol*. 2017 Jan 1;66(1):e47–65.
149. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. *PLoS Comput Biol*. 2009 Sep 25;5(9):e1000520.
150. IBGE | Cidades@ | São Paulo | São Caetano do Sul | Panorama [Internet]. [cited 2022 Jun 29]. Disponível em: <https://cidades.ibge.gov.br/brasil/sp/sao-caetano-do-sul/panorama>
151. São Caetano do Sul (SP) | Cidades e Estados | IBGE [Internet]. [cited

- 2023 Feb 11. Disponível em: <https://www.ibge.gov.br/cidades-e-estados/sp/sao-caetano-do-sul.html>.
152. Leal FE, Mendes-Correa MC, Buss LF, Costa SF, Bizario JCS, de Souza SRP, et al. Clinical features and natural history of the first 2073 suspected COVID-19 cases in the Corona São Caetano primary care programme: a prospective cohort study. *BMJ Open*. 2021 Jan 12;11(1):e042745.
 153. Artic Network [Internet]. [cited 2022 Jun 29]. Disponível em: <https://artic.network/ncov-2019>
 154. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc*. 2017 Jun;12(6):1261–76.
 155. Tyson JR, James P, Stoddart D, Sparks N, Wickenhagen A, Hall G, et al. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *BioRxiv*. 2020 Sep 4;
 156. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018 Sep 15;34(18):3094–100.
 157. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug 15;25(16):2078–9.
 158. O’Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol*. 2021 Jul 30;7(2):veab064.
 159. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol*. 2015 May 26;1(1):vev003.
 160. Fonseca PLC, Moreira FRR, de Souza RM, Guimarães NR, Carvalho NO, Adelino TER, et al. Tracking the turnover of SARS-CoV-2 VOCs Gamma to Delta in a Brazilian state (Minas Gerais) with a high-vaccination status. *Virus Evol*. 2022 Jul 27;8(2):veac064.
 161. Jackson B. Gofasta: command-line utilities for genomic epidemiology research. *Bioinformatics*. 2022 Aug 10;38(16):4033–5.
 162. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*. 2014 Nov 15;30(22):3276–8.
 163. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020 May 1;37(5):1530–4.

164. Yang Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol.* 1994 Sep;39(3):306–14.
165. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 2010 May;59(3):307–21.
166. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* 2016 Jan;2(1):vew007.
167. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 2018 Jan;4(1):vey016.
168. Didelot X, Siveroni I, Volz EM. Additive uncorrelated relaxed clock models for the dating of genomic epidemiology phylogenies. *Mol Biol Evol.* 2021 Jan 4;38(1):307–17.
169. Ghafari M, du Plessis L, Raghwani J, Bhatt S, Xu B, Pybus OG, et al. Purifying Selection Determines the Short-Term Time Dependency of Evolutionary Rates in SARS-CoV-2 and pH1N1 Influenza. *Mol Biol Evol.* 2022 Feb 3;39(2).
170. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol.* 2018 Sep 1;67(5):901–4.
171. O'Brien JD, Minin VN, Suchard MA. Learning to count: robust estimates for labeled distances between molecular sequences. *Mol Biol Evol.* 2009 Apr;26(4):801–14.
172. Minin VN, Suchard MA. Fast, accurate and simulation-free stochastic mapping. *Philos Trans R Soc Lond B Biol Sci.* 2008 Dec 27;363(1512):3985–95.
173. Buss L, Prete CA, Whittaker C, Salomon T, Oikawa MK, Pereira RHM, et al. Predicting SARS-CoV-2 Variant Spread in a Completely Seropositive Population Using Semi-Quantitative Antibody Measurements in Blood Donors. *Vaccines (Basel).* 2022 Aug 31;10(9).
174. Parag KV. Improved estimation of time-varying reproduction numbers at low case incidence and between epidemic waves. *PLoS Comput Biol.* 2021 Sep 7;17(9):e1009347.
175. Prete CA, Buss L, Dighe A, Porto VB, da Silva Candido D, Ghilardi F, et al. Serial interval distribution of SARS-CoV-2 infection in Brazil. *J Travel Med.* 2021 Feb 23;28(2).
176. Prete CA, Buss LF, Whittaker C, Salomon T, Oikawa MK, Pereira RHM,

- et al. SARS-CoV-2 antibody dynamics in blood donors and COVID-19 epidemiology in eight Brazilian state capitals: A serial cross-sectional study. *eLife*. 2022 Sep 22;11.
177. Brizzi A, Whittaker C, Servo LMS, Hawryluk I, Prete CA, de Souza WM, et al. Spatial and temporal fluctuations in COVID-19 fatality rates in Brazilian hospitals. *Nat Med*. 2022 Jul;28(7):1476–85.
 178. Banho CA, Sacchetto L, Campos GRF, Bittar C, Possebon FS, Ullmann LS, et al. Impact of SARS-CoV-2 Gamma lineage introduction and COVID-19 vaccination on the epidemiological landscape of a Brazilian city. *Commun Med (London)*. 2022 Apr 13;2:41.
 179. Luna-Muschi A, Borges IC, de Faria E, Barboza AS, Maia FL, Leme MD, et al. Clinical features of COVID-19 by SARS-CoV-2 Gamma variant: A prospective cohort study of vaccinated and unvaccinated healthcare workers. *J Infect*. 2022 Feb;84(2):248–88.
 180. Brito AF, Semenova E, Dudas G, Hassler GW, Kalinich CC, Kraemer MUG, et al. Global disparities in SARS-CoV-2 genomic surveillance. *Nat Commun*. 2022 Nov 16;13(1):7003.
 181. End-to-end integration of SARS-CoV-2 and influenza sentinel surveillance: revised interim guidance [Internet]. [cited 2023 Mar 16]. Disponivel em: https://www.who.int/publications/i/item/WHO-2019-nCoV-Integrated_sentinel_surveillance-2022.1
 182. Lemey P, Hong SL, Hill V, Baele G, Poletto C, Colizza V, et al. Accommodating individual travel history and unsampled diversity in Bayesian phylogeographic inference of SARS-CoV-2. *Nat Commun*. 2020 Oct 9;11(1):5110.
 183. Inward RPD, Parag KV, Faria NR. Using multiple sampling strategies to estimate SARS-CoV-2 epidemiological parameters from genomic sequencing data. *Nat Commun*. 2022 Sep 23;13(1):5587.