Brazilian Journal of Biology®

ISSN 1519-6984 (Print)
ISSN 1678-4375 (Online)

Original Article

Characterization of SARS-CoV-2 variants associated with reinfection cases in Botucatu, Brazil: a comparative genomic study

Caracterização das variantes do SARS-CoV-2 associadas a casos de reinfecção em Botucatu, Brasil: um estudo genômico comparativo

A. M. S. Kubo[a], R. A. Silveira[b], D. B. S. Figueiredo[b], L. A. Umezawa[b], K. I. Tasca[c], C. M. C. B. Fortaleza[c], G. T. Valente[b#] and R. M. T. Grotto[a,b#*]

[a]Universidade Estadual Paulista – UNESP, Faculdade de Ciências Agronômicas, Botucatu, SP, Brasil

[b]Hospital das Clínicas da Faculdade de Medicina de Botucatu da Universidade Estadual Paulista – HCFMB, Laboratório de Biotecnologia Aplicada, Laboratório de Biologia Molecular, Botucatu, SP, Brasil

[c]Universidade Estadual Paulista – UNESP, Faculdade de Medicina de Botucatu, Botucatu, SP, Brasil

[#]These authors contributed equally to this paper.

Abstract

The global spread of SARS-CoV-2 triggered the COVID-19 pandemic, resulting in a significant number of deaths. Variants such as B.1.1.7, B.1.351, P.1, and BA.1 distinguished by spike protein mutations arose due to immune pressure and lack of repair mechanisms. These changes amplified infectivity, increasing the chances of reinfection. We retrospectively included patients who have two positive tests for SARS-CoV-2 detected by real-time reverse transcription polymerase chain reaction with an interval of 90 days, or more, between 01 March 2020 and 30 April 2022. SARS-CoV-2 whole genome sequencing was performed for both infections, and the variants were identified using the Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin) database. This study identified a total of 254 cases of reinfections, 50.8% of the cases caused by the P.1 and BA.1 variants. Compared to the state scenario, a lower detection of the B.1.617.2 variant was identified, highlighting the importance of mass vaccination in the city of Botucatu, Brazil. However, most second infections (98.8%) were identified during the BA.1 wave, showing the effect of viral adaptative mechanisms on infection dynamics even in a vaccinated population. Taken together, our results emphasize the importance of constant monitoring and the maintenance of COVID-19 prevention and control strategies.

Keywords: COVID-19, reinfection, P.1, BA.1, genomic and phylogenetic analysis.

Resumo

A disseminação global do SARS-CoV-2 desencadeou a pandemia de COVID-19, resultando em um número significativo de mortes. Variantes como B.1.1.7, B.1.351, P.1 e BA.1, caracterizadas por mutações na proteína spike, surgiram devido à pressão imunológica e à falta de mecanismos de reparo. Essas alterações amplificaram a infectividade, aumentando as chances de reinfecção. Incluímos, retrospectivamente, pacientes com dois testes positivos para SARS-CoV-2, detectados por reação em cadeia da polimerase com transcrição reversa em tempo real, com intervalo de 90 dias ou mais, entre 01 de março de 2020 e 30 de abril de 2022. O sequenciamento do genoma completo do SARS-CoV-2 foi realizado para ambas as infecções, e as variantes foram identificadas utilizando o banco de dados Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin). Este estudo identificou um total de 254 casos de reinfecções, com 50,8% dos casos causados pelas variantes P.1 e BA.1. Comparado ao cenário estadual, uma menor detecção da variante B.1.617.2 foi observada, destacando a importância da vacinação em massa na cidade de Botucatu, Brasil. No entanto, a maioria das segundas infecções (98,8%) foi identificada durante a onda da BA.1, evidenciando o efeito dos mecanismos adaptativos virais na dinâmica de infecção, mesmo em uma população vacinada. Em conjunto, nossos resultados enfatizam a importância do monitoramento constante e da manutenção das estratégias de prevenção e controle da COVID-19.

Palavras-chave: COVID-19, reinfecção, P.1, BA.1, análise genômica e filogenética.

1. Introduction

The severe acute respiratory syndrome caused by SARS-CoV-2 spread exponentially throughout the world, and the coronavirus disease 2019 (COVID-19) was declared a global pandemic on 11 March 2020. According to the World Health Organization, COVID-19 caused more than 6.9 million deaths worldwide until September 21, 2023 (WHO, 2024), with Brazil accounting for 10% of these fatalities.

*e-mail: rejane.grotto@unesp.br
Received: November 18, 2024 – Accepted: March 17, 2025

The lack of a viral genetic repair mechanism and the evolutionary pressure exerted by host immune cell defense mechanisms (Volz et al., 2021) are part of the processes involved in the emergence of the new variants of SARS-CoV-2 (The Lancet, 2021; Zhao et al., 2022). Mutations in these variants have been described mainly in the S protein (Chen et al., 2020; The Lancet, 2021), a structural protein that mediates infection of host cells (The Lancet, 2021; Volz et al., 2021; Zhao et al., 2022). S protein S1 subunit contains a receptor-binding domain (RBD) that binds to the angiotensin-converting enzyme 2 (ACE2) receptor on the surface of human cells. The S2 subunit facilitates the fusion of the virus membrane with the host(Benetti et al., 2020; Letko et al., 2020; Zhou et al., 2021).

The continuous emergence of new variants is directly implicated in the spread of infection and in an increase in cases of COVID-19 reinfection (He et al., 2021).

The confirmation of SARS-CoV-2 reinfection requires a combination of diagnostic and genomic techniques (Yahav et al., 2021). Two positive polymerase chain reaction **(**PCR) tests with cycle threshold (Ct) values lower than 35 provide initial evidence of reinfection. The European Centre for Disease Prevention and Control (ECDC) advocates for a symptom-free interval between episodes along with a negative test and suggested Whole Genome Sequencing (WGS) to provide definitive confirmation by demonstrating that the two infections were caused by distinct viral strains (ECDC, 2020).Whereas the Centers for Disease Control and Prevention (CDC) suggests a minimum interval of 90 days between infections to substantiate the confirmation of SARS-CoV-2 reinfection (ECDC, 2020; CDC, 2019).

Currently, only a limited number of studies in Brazil have performed whole SARS-CoV-2 genome sequence for assessing possible reinfection cases (Santos et al., 2021; Penetra et al., 2023). In this context, we aimed in this study to investigate reinfection cases in the region of Botucatu, state of Sao Paulo by a next generation sequencing (NGS) methodology. Moreover, a phylogenetic analysis allowed us to compare the SARS-CoV-2 variants identified in our study to that reported for the state context. The study analyzed 68,613 positive cases of SARS-CoV-2, identifying 429 possible reinfections, and subsequently confirming 254 reinfection cases, with the P.1 variant identified as the cause of the initial infection and the BA.1 variant as the cause of the subsequent reinfection.

2. Materials and Methods

2.1. Eligibility criteria

The study encompassed retrospective samples obtained from residents of 11 cities within the Botucatu region. These samples were tested for SARS-CoV-2 between March 1, 2020, and April 30, 2022, at the Laboratory of Applied Biotechnology, Clinical Hospital of Botucatu Medical School (HCFMB) in Botucatu, Brazil.

The inclusion criteria involved individuals with two positive tests for SARS-Cov-2 infection detected by real-time reverse transcription polymerase chain reaction (RT-qPCR) assay and collected at an interval of ≥ 90 days. Samples that exhibited the same strain in genomic sequencing in the first and second samples were excluded.

This study was approved by the Institutional Ethics Committee of the Faculty of Medicine of Botucatu (FMB) (Process CAAE: 57919122.9.0000.5411).

2.2. Real-Time RT-qPCR for SARS-CoV-2 detection

Specimens of nasopharyngeal and oropharyngeal swabs collected from patients were suspended in tubes containing 0.9% saline media and transported to the Laboratory of Applied Biotechnology. After vortexing, 250 µL were portioned for diagnostic purposes, while 1.5 mL were set aside as a backup. The RNA extraction was performed utilizing a magnetic bead-based solid-phase reversible immobilization method using an automated platform (Extract 96, Loc-cus, Brazil).

RT-qPCR assays were performed using the GeneFinderTM COVID-19 Fast RealAmp kit (Osang HealthCare, Republic of Korea). The CFX96 Touch Real-Time Detection System (BioRad, Hercules, California, USA) and the Applied BiosystemsTM 7500 Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA) were used for amplification and detection of SARS-CoV-2. The RT-qPCR cycling conditions used on the CFX96 Touch Real-Time Detection System were 5 minutes at 50 °C; 20 seconds at 95 °C; and 45 cycles of 95 °C for 1 second and 58 °C for 5 seconds. The following amplification conditions were performed on the Applied BiosystemsTM 7500 Real-Time PCR System: 50 °C for 20 minutes; 95 °C for 5 minutes; and 45 cycles of 95 °C for 15 seconds and 58 °C for 1 minute.

The assay for detection of SARS-CoV-2 targeted three viral genes: envelope gene (E gene), nucleocapsid protein gene (N gene), and RNA-dependent RNA polymerase (RdRp gene). We considered a sample positive for SARS-Cov-2 if amplification was detected with a Ct value <40 for at least two viral genes. Additionally, the RNaseP gene was used as an internal control, considering eligible for analysis, samples with a Ct value <35. All test and other analysis procedures were performed following the manufacturer's instructions.

2.3. SARS-CoV-2 whole genome sequencing

2.3.1. Ion AmpliSeq™ SARS CoV 2 Insight Research Assay

The viral RNA was reverse transcribed using the Ion Torrent™ NGS Reverse Transcription kit following the manufacturer's instructions. The RT reaction was performed using the T100 BioRad termocycler (BioRad, Hercules, California, USA) at 25 °C for 10 minutes, followed by 50 °C for 10 minutes and 85 °C for 5 minutes.

Sequencing libraries were prepared using the Ion AmpliSeq™ Library Kit Plus and the Ion AmpliSeq SARS-Cov-2 Insight Research Panel according to the manufacturer's instructions. The panel consists of two primer pools with amplicons ranging from 125 bp to 275 bp in length, covering 99% of the SARS-CoV-2 genome, allowing identification of all variants (Thermo Fisher Scientific, 2023).

The DNA target amplification reaction was prepared using 4.5 µL of 5X Ion AmpliSeq™ HiFi Mix, 10 µL of cDNA, and 3.5 µL of nuclease-free water. For each sample, two

separated reactions were performed containing 8 µL of the previous described mastermix and 2 µL of the 5X Ion AmpliSeq™ Primer Pool 1 or 5X Ion AmpliSeq™ Primer Pool 2. Thermal cycling was carried out under the following conditions: 98 °C for 2 minutes, and 20 cycles at 98 °C for 15 seconds and 60 °C for 4 minutes.

After amplification of target regions, the two sets of amplification reactions were combined and partially digested with 2 µL of FuPa Reagent at 50 °C for 10 minutes, 55 °C for 10 minutes and 60 °C for 20 minutes.

Adapters were ligated to the amplified products using 4 µL of Switch Solution, 2 µL of diluted Ion Xpress™ Barcode and 2 µL of DNA Ligase. The samples were incubated at 22 °C for 30 minutes, 68 °C for 5 minutes, and 72 °C for 5 minutes. After barcoding, sets of 8 samples were pooled for the next steps.

The purification of the pooled unamplified libraries was performed using Agencourt™ AMPure™ XPReagent (Beckman Coulter).

Each pool of 8 samples was quantified by real time PCR using the Ion Library TaqMan™ Quantitation Kit. The pooled samples were diluted to 1:1000 and 1:10000 and a set of four serial dilutions was prepared for the control library (*Escherichia coli* DH10B). The reaction mixture was prepared using 5 µL of 2X Ion Library qPCR Master Mix, 0.5µL of Ion Library TaqMan™ Quantitation Assay 20X, and 4.5 µL of each pooled sample or library control. Real-time PCR quantification was performed using the QuantStudio™3 Real-Time PCR System (ThermoFisher, Waltham, MA, USA) under the following conditions: 50 °C for 2 minutes, 95 °C for 2 minutes, and 40 cycles of 95 °C for 15 seconds and 60 °C for 1 minute.

After quantification, each pool was diluted and combined to achieve a 30 pM or a 50 pM solution, for use with the Ion 530™ Chip or Ion 540™ Chip, respectively.

Clonal amplification by emulsion PCR was performed on the Ion Torrent Ion Chef System. We used either the Ion 530 Chip or the Ion 540 Chip, depending on the number of samples available per run. Whole genome sequencing was performed using the IonGeneStudio™ S5 System.

2.3.2. Illumina COVIDSeq test

Thirteen samples were processed in accordance with the manufacturer's instructions for Illumina COVIDSeq.

The viral RNA underwent annealing using random hexamers to facilitate cDNA synthesis, employing 8.5 µL of Elution Prime Fragment 3HC Mix HT according to the manufacturer's guidelines. The reverse transcription (RT) process was conducted using the T100 BioRad thermocycler (BioRad, Hercules, California, USA) at 65 °C for 3 minutes, followed by a hold at 4 °C.

The synthesis of cDNA was executed using 9 µL of First Strand Mix H and 1 µL of Reverse Transcriptase HT for each sample. The RT reaction was carried out in a thermocycler at 25 °C for 5 minutes, followed by 50 °C for 10 minutes and 85 °C for 5 minutes.

The amplification of the DNA target involved preparing a reaction mixture comprising 15 µL of Illumina PCR Mix HT, 3.5 µL of nuclease-free water, and 4.3 µL of either COVIDSeq Primer Pool 1 HT or COVIDSeq Primer Pool 2 HT. Two separate reactions were conducted for each sample, each containing 20 µL of the master mix along with either COVIDSeq Primer Pool 1 HT or COVIDSeq Primer Pool 2 HT, and 5 µL of cDNA. Thermal cycling was carried out under the following conditions: 98 °C for 3 minutes, followed by 35 cycles at 98 °C for 15 seconds and 63 °C for 5 minutes.

Following the amplification of target regions, 10 µL from both reactions were combined with 30 µL of a master mix comprising 4 µL of Enrichment BLT HT, 12 µL of Tagmentation Buffer 1 H, and 20 µL of Nuclease-free water. Cycling was conducted at 55 °C for 5 minutes.

The purification of adapter-tagged amplicons was performed using of Stop Tagment Buffer 2 HT as per the manufacturer's guidelines.

The Amplify Tagmented Amplicons step proceeded using 10 µl index adapters and 40 µL of the master mix composed of 24 µL Enhanced PCR Mix H and 24 µL of nuclease-free water. The samples were incubated at 72 °C for 3 minutes, 98 °C for 3 minutes, 7 cycles of 98 °C for 20 seconds, 60 °C for 30 seconds and 72 °C for 1 minute, and finally 72 °C for 3 minutes.

After amplification, sets of 8 samples were combined into pools for subsequent steps. Then, 55 µL from each pool was transferred into a 1.5 ml tube along with 396 µL of Illumina Tune Beads. The libraries were then purified using freshly prepared 80% ethanol following the instructions outlined in the manual (Illumina, 2023).

The libraries were quantified using 2 µl library pool into a Qubit dsDNA HS Assay kit, and normalized at 4 nM using Resuspension Buffer HT.

After normalized, the library was denatured with NaOH (0.2N) and normalized at 7 pM. Whole genome sequencing was performed using the flow cell V2 in a MiSeq equipment.

2.4. Variants identification

The FASTQ files obtained were imported into Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin) database online: this tool allows tracking the transmission and spread of SARS-CoV-2 identifying variants by sequence data analyses (Rambaut et al., 2020a). This database aligns the query sequences against the reference SARS-CoV-2 sequence. Then, using minimap2(Li, 2018), the tool creates a sequence alignment and generates a FASTA file with gofasta (Jackson, 2022) that contains non-coding regions masked with N's. This file was streamlined into unique sequences using the hash system and processed through the designation cache to identify a designated lineage in the Pangolin database. A sequence QC check was performed, which reported the percentage of ambiguities in sequences. Scorpio (Cov-Lineages, 2023) analyzed variants defined by the constellation of the Cov lineage. The output of all steps is compiled into a report CSV file. Then, the interference pipelines passed by either pangoLEARN or UShER and processed into a standardized format. The conclusive report comprised variant pipelines.

For phylogenetic inference, redundancies among all sequences were excluded using the CDHit (Fu et al., 2012) (-c 0.99, -M 0, -AS, -AL): -c is sequence identity threshold, -M is use max available memory, -AS is alignment coverage

control for the shorter sequence, and -AL is alignment coverage control for the longer sequence.The 82 remaining sequences including the Wuhan sequence (hCoV19/Wuhan/WIV04/2019|EPI_ISL_402124) were aligned using the offline Clustal Omega (v.1.2.2) implemented in Seaview (Gouy et al., 2010). The misaligned portions were trimmed using the TrimAI on XSEDe implemented in the CIPRES Science Gateway(Miller et al., 2010). Some gaps were manualy excluded using the Geneious (v.2023.2.1) software. The best evolutionary model was inferred using the Jmodel test (Posada, 2008) 7 substitution schemes, I+F+G, 4 categories, base tree for likelihood calculation = ML optimized, base tree search = best). We also performed the BIC result analysis (default) in the Jmodel test. The phylogenetic tree was inferred using the PhyML (with aLRT branch support, GTR+I evolutionary model, with I=0.9170, best of NNI and SPR, and BioNJ starting tree) implemented in the Seaview(Gouy et al., 2021). The tree was visualized and edited using the FigTree software. The Wuhan sequence was used as an outgroup.

Epidemiological data on COVID-19 infections in the State of São Paulo were collected from the Genomic Network Fiocruz Dashboard (FIOCRUZ, 2023) from March 1, 2020, to April 30, 2022.

Comparative analysis with SARS-CoV-2 variants in circulation in the State of São Paulo from March 1, 2020 to April 30, 2022 was performed accessing FioCruz´s Genomic Network (FIOCRUZ, 2023).

2.5. Reinfection criteria

The study implemented rigorous reinfection criteria that aligned with both the European CDC and the CDC guidelines. Only samples exhibiting an interval of at least 90 days between positive SARS-CoV-2 diagnoses acquired by RT-qPCR assays were considered. Subsequently, a comprehensive analysis of complete virus sequencing by NGS was conducted to corroborate genetic differentiation in infected patients.

3. Results

3.1. Reinfection cases

We identified 254 cases of reinfections by SARS-CoV-2 in patients from the region of Botucatu, São Paulo State, Brazil. P.1 and BA.1 lineages were prevalent in the first and second infections, respectively, with 129 and 252 cases (Table 1). Most first infections occur between 2020 and 2021 (99.61%), while most of second infections occurred at 2022 (99.61%).

In the first infection episode, the most prevalent variants were P.2 and P.1, with 24.02% (61 cases) and 50.79% (129 cases), respectively (Figure 1). Only B.1.1, P.1, and BA.1 were identified in the second infection, in which BA.1 was ranked as the most prevalent variant (98.82%, accounting for 252 cases) (Figure 2).

3.2. Phylogenetic analysis of SARS-CoV-2 variants

Phylogenetic tree analysis grouped all variants in single well supported branches, being the variants B.1.1 and P2 the only exceptions. The variants BA.2, BA.1 and B.1.1.7 formed a very well supported clade. Some P.2 and B.1.1 variants were dispersed in different clades (see the right-side arrows in Figure 3), although P.2 variants form a large clade. In this case, the P.2 sequence 14017 is a sister group of B.1.1, whereas the sequence 10540 B.1.1 variant is a sister group of B.1.1.28 variants (Figure 3).

Table 1. Variants involved in reinfection cases by SARS-CoV-2 at Botucatu, Sao Paulo State, Brazil.

	Second infection		
First infection	**B.1.1**	**P.1**	**BA.1**
B.1.1	0	1	29
B.1.1.28	0	0	10
B.1.1.33	0	0	2
P.2	1	0	60
B.1.1.7	0	0	10
B.1.617.2	0	0	6
P.1	0	0	129
BA.1	0	1	0
Others	0	0	5

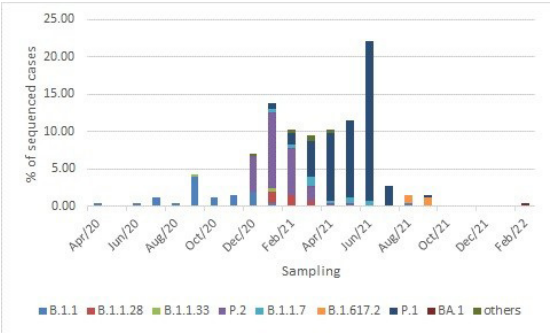

Figure 1. Temporal distribution of SARS-CoV-2 lineages in first infection.

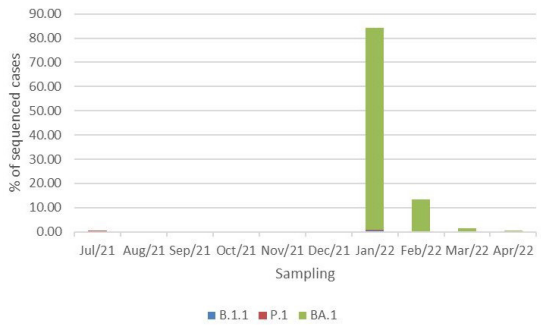

Figure 2. Temporal distribution of SARS-CoV-2 lineages in second infection.

For statistical analysis, all subvariants of BA.1 and BA.2 (clustered yellow and purple) was clustered into BA.1 lineage, B.1.1.7 variants (clustered orange) into B.1.1.7 lineage, P.1. variants (clustered light blue) into P.1 lineage, B.1.1 variants (clustered dark blue) into B.1.1 lineage, B.1.1.28 variants (clustered cyan blue) into B.1.1.28 lineage, P.2. variants (clustered green) into P.2 lineage (Figure 3). The other variants were grouped according to the Fiocruz´s Genomic Network (Table 2).

According to Fiocruz´s Genomic Network, from March 1, 2020 to April 30, 2022 the most prevalent variants in the State of São Paulo were P.1 (40.78%), B.1.617.2 (30.60%) e BA.1 (22.96%) (Figure 4). In the same period, we observed an elevated incidence of P.1 (25.79%) and BA.1 (49.61%) variants. Notably, a lower prevalence (1.18%) of B.1.617.2 was observed in this study.

4. Discussion and Conclusions

We confirmed 254 SARS-CoV-2 reinfection cases in the region of Botucatu, São Paulo State, Brazil, by whole genome sequencing using next generation sequencing technology. Furthermore, the clustering of SARS-CoV-2 variants by a phylogenetic analysis allowed us to compare our data with the epidemiological profile of circulating variants in the Sao Paulo State.

The BA.2 and BA.1 sub-lineages of the Omicron variant share 38 mutations, with the majority (27) located in the spike receptor-binding domain (RBD) (COV-LINEAGES, 2021), accounting for the observed phylogenetic closeness. Furthermore, both variants, along with B.1.1.7, exhibit mutations N501Y and P681H, which contribute to the clustering evident in the phylogenetic tree (Rambaut et al., 2020b; COV-LINEAGES, 2021).

The dispersion of the P.2 variant in various clades can be attributed to its lineage traced back to the ancestral B.1.1.28 strain(Lamarca et al., 2021), which was the precursor to the B.1 lineage (Hill et al., 2022).

The global spread of SARS-CoV-2 has been marked by successive waves driven by dominant variants. The first surge, between late 2020 and early 2021, was largely attributed to the B.1.1.7 and B.1.351 variants, leading to a sharp rise in infections and deaths. In the following months, the P.1 variant caused another peak, particularly in April and May 2021. By August 2021, the highly transmissible B. 1.617. 2

Table 2. Phylogenetic classification of variants identified in reinfection cases.

Variants	Lineage
B.1.1.7	B.1.1.7
B.1.1	B.1.1
B.1.1.28	B.1.1.28
B.1.1.33	B.1.1.33
AY.99.2	B.1.617.2
B.1.617.2	B.1.617.2
P.1	P.1
P.1.2	P.1
P.1.7	P.1
P.1.14	P.1
BA.1.1	BA.1
BA.1.14	BA.1
BA.1.14.1	BA.1
BA.1.14.2	BA.1
BA.1.15	BA.1
BA.1.9	BA.1
BA.2	BA.1
BA.2.9	BA.1
B.1	Others
B.1.1.376	Others
N.9	Others
P.2	P.2

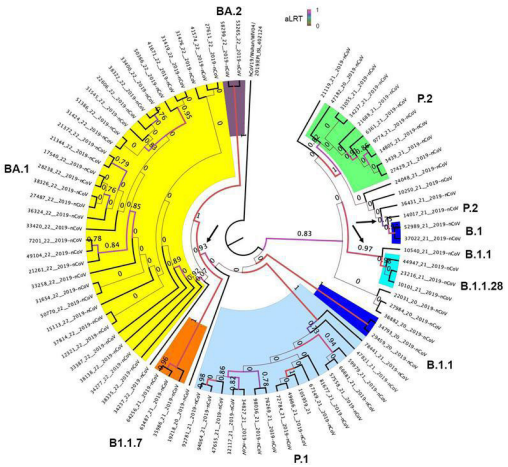

Figure 3. Cladogram inferred by maximum likelihood. The aLRT (approximate likelihood-ration test) was inferred as the branch support. The names of variants are depicted surrounding the cladogram. The arrows evidenced well supported branches that group different variants.

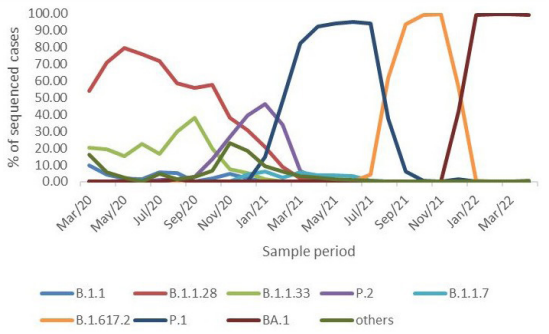

Figure *4.* Temporal distribution of SARS-CoV-2 lineages in State of São Paulo. Source: Adapted from Fiocruz Genomic Network.

variant fueled a new wave of cases and fatalities. Later, in December 2021 and January 2022, the Omicron variant led to an unprecedented spike in global infections due to its high transmissibility and ability to evade immune responses. In March 2022, the BA.2 subvariant quickly overtook BA.1 as the predominant strain, followed by BA.4 and BA.5, which became dominant around July 2022, further shaping the course of the pandemic (Chen et al., 2023; Thakur et al., 2022)

Similarly, during the study period, Brazil experienced three distinct SARS-CoV-2 waves.

The first between February 2020 to November 2020 with the presence of multiple variants, the second between December 2020 to December 2021 with P.2 and P.1 co-dominance, and finally the third in early 2022 with the prevalence of Omicron (Alcantara et al., 2022; Martins-Filho et al., 2022).

The P.1 variant was first detected in northern Brazil at the end of 2020 and quickly spread throughout the country (Banho et al., 2022; Spira, 2022), peaking in the Botucatu-SP region in June 2021. It has 10 non-synonymous mutations in the S protein, having important mutations in the spike RBD: K417T, N501Y and E484K (Naveca et al., 2021; Zimerman et al., 2022). The E484K mutation is associated with escape from neutralizing antibodies, increasing the possibility of reinfection (Nelson et al., 2021; Faria et al., 2021; Zimerman et al., 2022), while the N501Y mutation increased binding specific and rapid lineage growth (Faria et al., 2021; Zimerman et al., 2022). The K417T mutation can diminish the activity of some monoclonal antibodies (Zimerman et al., 2022). All these characteristics resulted in a more infectious variant with immune escape mechanisms. This accounts for the majority of the first infections observed in this study.

An increased number of second infections was observed during the outbreak of SARS-CoV-2 BA.1 variant that occurred mostly in the beginning of 2022. In the same period, 99.33% of all sequences deposited from the Sao Paulo State were attributed to the BA.1 variant. Both occurrences can be explained by the fact that the BA.1 lineage is the variant with the largest number of mutation sites among the SARS-CoV-2 variants characterized to date (CoVariantes, 2023). This variant has 30 substitutions, 3 deletions, and 1 insertion in the S protein (Naveca et al., 2023), leading to a greater improvement in transmission and immune evasion mechanisms. For instance, the Q498R mutation enhances the binding affinity of the S-protein to ACE2, and E484A is associated with immune escape due to the fact that this residue affects the antibody binding (Naveca et al., 2023; Nelson et al., 2021). This supports the increased occurrence of reinfections with this variant, even in 3 doses of spike-based vaccinated people.

It is noteworthy that, according to the eligibility criteria established for cases of SARS-CoV-2 reinfection adopted by this study, instances involving the BA.1 variant as the initial strain of infection may not have been detected. The Botucatu city, in the interior of Sao Paulo State, Brazil, was selected to participate in a large-scale study to monitor the safety of the ChAdOx1-nCoV19 vaccine. This vaccine was developed by the University of Oxford and AstraZeneca, produced and distributed in Brazil by the Oswaldo Cruz Foundation (FioCruz). The mass vaccination coverage of the eligible adult population was 84.2% and 80.2% for the first and second doses, respectively. The two doses were administered between May 16, 2021 and the beginning of September 2021 (Clemens et al., 2022). A total of 77.683 adults were vaccinated and 151.734 doses administrated. In Botucatu, the population between 18 and 60 years of age, except for healthcare workers and high-risk individuals, received the ChAdOx1-nCoV19 vaccine exclusively (Costa Clemens et al., 2022) .The discrepancy observed in the detection of the B.1.617.2 variant between our study and the state of Sao Paulo can be attributed to the substantial reduction in cases following the second dose of ChAdOx1-nCoV19 vaccine. Only six cases, related to cases of reinfection, were detected in Botucatu during the period immediately after the second dose of the vaccine, highlighting the potential protective effect of vaccination against some variants.

However, the reinfection profile found here corroborates the concerns surrounding immune evasion by variants, particularly BA.1, which exhibits a distinct mutational pattern. The ability of the BA.1 variant to partially evade previously acquired immunity underscores the imperative for constant surveillance. This further emphasizes the importance of continually adapting COVID-19 prevention and control strategies.

Acknowledgements

We extend our sincere appreciation to all the members of the Molecular Biology Laboratory for their invaluable assistance in the diagnostics and analysis of the samples. Special thanks to Rita de Cássia Alvarado for her exceptional technical support. The authors gratefully acknowledge UNESP for its institutional support. This research was funded by grant 401575/2021-7, National Council for Scientific and Technological Development (CNPq), grant 2021/13767-4, São Paulo Research Foundation (FAPESP) and grant 2022/02261-5, São Paulo Research Foundation (FAPESP).

References

ALCANTARA, L.C.J., NOGUEIRA, E., SHUAB, G., TOSTA, S., FRISTCH, H., PIMENTEL, V., SOUZA-NETO, J.A., COUTINHO, L.L., FUKUMASU, H., SAMPAIO, S.C., ELIAS, M.C., KASHIMA, S., SLAVOV, S.N., CICCOZZI, M., CELLA, E., LOURENCO, J., FONSECA, V. and GIOVANETTI, M., 2022. SARS-CoV-2 epidemic in Brazil: how the displacement of variants has driven distinct epidemic waves. *Virus Research*, vol. 315, pp. 198785. http://doi.org/10.1016/j.virusres.2022.198785. PMid:35461905.

BANHO, C.A., SACCHETTO, L., CAMPOS, G.R.F., BITTAR, C., POSSEBON, F.S., ULLMANN, L.S., MARQUES, B.C., DA SILVA, G.C.D., MORAES, M.M., PARRA, M.C.P., NEGRI, A.F., BOLDRIN, A.C., BARCELOS, M.D., DOS SANTOS, T.M.I.L., MILHIM, B.H.G.A., ROCHA, L.C., DOURADO, F.S., DOS SANTOS, A.L., CICONI, V.B., PATUTO, C., VERSIANI, A.F., DA SILVA, R.A., DE OLIVEIRA LOBL, E.E., HERNANDES, V.M., ZINI, N., PACCA, C.C., ESTOFOLETE, C.F., FERREIRA, H.L., RAHAL, P., ARAÚJO JUNIOR, J.P., COHEN, J.A., KERR, C.C., ALTHOUSE, B.M., VASILAKIS, N. and NOGUEIRA, M.L., 2022. Impact of SARS-CoV-2 Gamma lineage introduction and COVID-19 vaccination on the epidemiological landscape of a Brazilian city. *Communication & Medicine*, vol. 2, no. 1, pp. 41. http://doi.org/10.1038/s43856-022-00108-5.

BENETTI, E., TITA, R., SPIGA, O., CIOLFI, A., BIROLO, G., BRUSELLES, A., DODDATO, G., GILIBERTI, A., MARCONI, C., MUSACCHIA, F., PIPPUCCI, T., TORELLA, A., TREZZA, A., VALENTINO, F., BALDASSARRI, M., BRUSCO, A., ASSELTA, R., BRUTTINI, M., FURINI, S., SERI, M., NIGRO, V., MATULLO, G., TARTAGLIA, M., MARI, F., FRULLANTI, E., FALLERINI, C., DAGA, S., CROCI, S., AMITRANO, S., FAVA, F., MONTAGNANI, F., DI SARNO, L., TOMMASI, A., PALMIERI, M., EMILIOZZI, A., FABBIANI, M., ROSSETTI, B., ZANELLI, G., BERGANTINI, L., D'ALESSANDRO, M., CAMELI, P., BENNET, D., ANEDDA, F., MARCANTONIO, S., SCOLLETTA, S., FRANCHI, F., MAZZEI, M.A., CONTICINI, E., CANTARINI, L., FREDIANI, B., TACCONI, D., FERI, M., SCALA, R., SPARGI, G., CORRIDI, M., NENCIONI, C., CALDARELLI, G.P., SPAGNESI, M., PIACENTINI, P., BANDINI, M., DESANCTIS, E., CANACCINI, A., SPERTILLI, C., DONATI, A., GUIDELLI, L., CROCI, L., VERZURI, A., ANEMOLI, V., OGNIBENE, A., VAGHI, M., D'ARMINIO MONFORTE, A., MERLINI, E., MONDELLI, M.U., MANTOVANI, S., LUDOVISI, S., GIRARDIS, M., VENTURELLI, S., SITA, M., COSSARIZZA, A., ANTINORI, A., VERGORI, A., RUSCONI, S., SIANO, M., GABRIELI, A., RIVA, A., FRANCISCI, D., SCHIAROLI, E., SCOTTON, P.G., ANDRETTA, F., PANESE, S., SCAGGIANTE, R., PARISI, S.G., CASTELLI, F., QUIROS-ROLDAN, M.E., MAGRO, P., MINARDI, C., CASTELLI, D., POLESINI, I., DELLA MONICA, M., PISCOPO, C., CAPASSO, M., RUSSO, R., ANDOLFO, I., IOLASCON, A., CARELLA, M., CASTORI, M., MERLA, G., AUCELLA, F., RAGGI, P., MARCIANO, C., PERNA, R., BASSETTI, M., DI BIAGIO, A., SANGUINETTI, M., MASUCCI, L., GABBI, C., VALENTE, S., GUERRINI, S., MELONI, I., MENCARELLI, M.A., RIZZO, C.L., BARGAGLI, E., MANDALÀ, M., GIORLI, A., SALERNI, L., FIORENTINO, G., ZUCCHI, P., PARRAVICINI, P., MENATTI, E., BARATTI, S., TROTTA, T., GIANNATTASIO, F., COIRO, G., LENA, F., COVIELLO, D.A., MUSSINI, C., RENIERI, A. and PINTO, A.M., 2020. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *European Journal of Human Genetics*, vol. 28, no. 11, pp. 1602-1614. http://doi.org/10.1038/s41431-020-0691-z. PMid:32681121.

CENTER FOR DISEASE CONTROL AND PREVENTION – CDC ., 2019 [viewed 27 September 2023]. *What is COVID-19 Reinfection?* [online]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/your-health/reinfection.html

CHEN, J., WANG, R., HOZUMI, Y., LIU, G., QIU, Y., WEI, X. and WEI, G.W., 2023. Emerging dominant SARS-CoV-2 variants. *Journal of Chemical Information and Modeling*, vol. 63, no. 1, pp. 335-342. http://doi.org/10.1021/acs.jcim.2c01352. PMid:36577010.

CHEN, Y., LIU, Q. and GUO, D., 2020. Emerging coronaviruses: genome structure, replication, and pathogenesis. *Journal of Medical Virology*, vol. 92, no. 10, pp. 2249. http://doi.org/10.1002/jmv.25681. PMid:31967327.

CLEMENS, S.A.C., FORTALEZA, C.M.C.B., CROWE, M., POLLARD, A., TASCA, K.I., GROTTO, R.M.T., MARTINS, M.R., SPADARO, A.G., BARRETTI, P., VERSTRAETEN, T. and CLEMENS, R., 2022. Safety of the Fiocruz ChAdOx COVID-19 vaccine used in a mass vaccination campaign in Botucatu, Brazil. *Vaccine*, vol. 40, no. 47, pp. 6722-6729. http://doi.org/10.1016/j.vaccine.2022.08.026. PMid:36055876.

COSTA CLEMENS, S.A., FORTALEZA, C.M.C.B., CROWE, M., TASCA, K.I., SPADARO, A.G., SOUZA-NETO, J.A., GROTTO, R.M.T., SIDER, R., JIMENO, J., VERSTRAETEN, T. and CLEMENS, R., 2022. Effectiveness of the Fiocruz recombinant ChadOx1-nCoV19 against variants of SARS-CoV-2 in the Municipality of Botucatu-SP. *Frontiers in Public Health*, vol. 10, pp. 1016402. http://doi.org/10.3389/fpubh.2022.1016402. PMid:36311567.

COVARIANTES, 2023 [viewed 25 February 2023]. CoVariants: 21K (Omicron) [online]. Available from: https://covariants.org/variants/21K.Omicron

COV-LINEAGES, 2021 [viewed 27 November 2023]. *Proposal to split B.1.1.529 to incorporate a newly characterised sibling lineage · Issue #361 · cov-lineages/pango-designation · GitHub* [software]. Available from: https://github.com/cov-lineages/pango-designation/issues/361

COV-LINEAGES, 2023 [viewed 25 February 2023]. *GitHub - cov-lineages/scorpio: serious constellations of reoccurring phylogenetically-independent origin* [software]. Available from: https://github.com/cov-lineages/scorpio

EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL – ECDC., 2020 [viewed 27 September 2023]. *Threat Assessment Brief: Reinfection with SARS-CoV-2: considerations for public health response* [online]. Available from: https://www.ecdc.europa.eu/en/publications-data/threat-assessment-brief-reinfection-sars-cov-2

FARIA, N.R., CLARO, I.M., CANDIDO, D., FRANCO, L.A.M., ANDRADE, P.S., COLETTI, T.M., SILVA, C.A.M., SALES, F.C., MANULI, E.R., AGUIAR, R.S., GABURO, N., CAMILO, C.C., FRAIJI, N.A., CRISPIM, M.A.E., CARVALHO, M.S.S., RAMBAUT, A., LOMAN, N., PYBUS, O.G., SABINO, E.C., 2021 [viewed 25 February 2023]. *Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings* [online]. Edinburgh: Virological. Available from: https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586

FU, L., NIU, B., ZHU, Z., WU, S. and LI, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, vol. 28, no. 23, pp. 3150-3152. http://doi.org/10.1093/bioinformatics/bts565. PMid:23060610.

FUNDAÇÃO OSWALDO CRUZ – FIOCRUZ, 2023 [viewed 25 February 2023]. *Dashboard-pt – Genomahcov – Fiocruz* [online]. Available from: https://www.genomahcov.fiocruz.br/dashboard-pt/

GOUY, M., GUINDON, S. and GASCUEL, O., 2010. SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, vol. 27, no. 2, pp. 221-224. http://doi.org/10.1093/molbev/msp259. PMid:19854763.

GOUY, M., TANNIER, E., COMTE, N. and PARSONS, D.P., 2021. Seaview Version 5: a multiplatform software for multiple sequence alignment, molecular phylogenetic analyses, and tree reconciliation. In: K. KATOH, ed. *Multiple sequence alignment*. New York, NY: Humana, pp. 241–260. Methods in Molecular Biology, vol. 2231. https://doi.org/10.1007/978-1-0716-1036-7_15.

HE, S., HAN, J. and LICHTFOUSE, E., 2021. Backward transmission of COVID-19 from humans to animals may propagate reinfections and induce vaccine failure. *Environmental Chemistry Letters*, vol. 19, no. 2, pp. 763-768. http://doi.org/10.1007/s10311-020-01140-4. PMid:33424524.

HILL, V., DU PLESSIS, L., PEACOCK, T.P., AGGARWAL, D., COLQUHOUN, R., CARABELLI, A.M., ELLABY, N., GALLAGHER, E., GROVES, N., JACKSON, B., MCCRONE, J.T., O'TOOLE, Á., PRICE, A., SANDERSON, T., SCHER, E., SOUTHGATE, J., VOLZ, E., BARCLAY, W.S., BARRETT, J.C., CHAND, M., CONNOR, T., GOODFELLOW, I., GUPTA, R.K., HARRISON, E.M., LOMAN, N., MYERS, R., ROBERTSON, D.L., PYBUS, O.G. and RAMBAUT, A., 2022. The origins and molecular evolution of SARS-CoV-2 lineage B.1.1.7 in the UK. *Virus Evolution*, vol. 8, no. 2, pp. veac080. http://doi.org/10.1093/ve/veac080. PMid:36533153.

ILLUMINA, 2023 [viewed 25 February 2023]. *Illumina CovidSeq Test Instructions for Use (# 1000000128490)* [online]. San Diego: Illumina, Inc. Available from: https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-

JACKSON, B., 2022. gofasta: command-line utilities for genomic epidemiology research. *Bioinformatics*, vol. 38, no. 16, pp. 4033-4035. http://doi.org/10.1093/bioinformatics/btac424. PMid:35789376.

LAMARCA, A.P., DE ALMEIDA, L.G.P., FRANCISCO JUNIOR, R.S., LIMA, L.F.A., SCORTECCI, K.C., PEREZ, V.P., BRUSTOLINI, O.J., SOUSA, E.S.S., SECCO, D.A., SANTOS, A.M.G., ALBUQUERQUE, G.R., MARIANO, A.P.M., MACIEL, B.M., GERBER, A.L., GUIMARÃES, A.P.C., NASCIMENTO, P.R., FREIRE NETO, F.P., GADELHA, S.R., PORTO, L.C., CAMPANA, E.H., JERONIMO, S.M.B. and VASCONCELOS, A.T.R., 2021. Genomic surveillance of SARS-CoV-2 tracks early interstate transmission of P.1 lineage and diversification within P.2 clade in Brazil. *PLoS Neglected Tropical Diseases*, vol. 15, no. 10, e0009835. http://doi.org/10.1371/journal.pntd.0009835. PMid:34644287.

LETKO, M., MARZI, A. and MUNSTER, V., 2020. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiology*, vol. 5, no. 4, pp. 562-569. http://doi.org/10.1038/s41564-020-0688-y. PMid:32094589.

LI, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, vol. 34, no. 18, pp. 3094-3100. http://doi.org/10.1093/bioinformatics/bty191. PMid:29750242.

MARTINS-FILHO, P.R., MARQUES, R.S., TAVARES, C.S.S., ARAÚJO, A.A.S. and QUINTANS-JÚNIOR, L.J., 2022. The Role of Vaccination in Preventing Deaths During the Omicron-driven Tsunami in Brazil. *Disaster Medicine and Public Health Preparedness*, vol. 16, no. 6, pp. 2252-2253. http://doi.org/10.1017/dmp.2022.146. PMid:35672263.

MILLER, M.A., PFEIFFER, W. and SCHWARTZ, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *2010 Gateway Computing Environments Workshop (GCE)*, 2010, New Orleans, LA, USA. New York: IEEE, pp. 1-8. http://doi.org/10.1109/GCE.2010.5676129

NAVECA, F., COSTA, C., NASCIMENTO, V., SOUZA, V., CORADO, A., NASCIMENTO, F., COSTA, A., DUARTE, D., SILVA, G., MEJÍA, M., PESSOA, K., GONÇALVES, L., BRANDÃO, M.J., JESUS, M., PINTO, R., SILVA, M., MATTOS, T., ABDALLA, L., SANTOS, J.H., COSTA-FILHO, R., WALLAU, G.L., SIQUEIRA, M.M., DELATORRE, E., GRÄF, T., BELLO, G., RESENDE, P.C., 2021 [viewed 25 February 2023]. *SARS-CoV-2 reinfection by the new Variant of Concern (VOC) P.1 in Amazonas, Brazil* [online]. Edinburgh: Virological. Available from: https://virological.org/t/sars-cov-2-reinfection-by-the-new-variant-of-concern-voc-p-1-in-amazonas-brazil/596

NAVECA, F.G., NASCIMENTO, V.A., NASCIMENTO, F., OGRZEWALSKA, M., PAUVOLID-CORRÊA, A., ARAÚJO, M.F., ARANTES, I., BATISTA, É.R., MAGALHÃES, A.Á., VINHAL, F., MATTOS, T.P., RIEDIGER, I., DEBUR, M., GRINSZTEJN, B., VELOSO, V.G., BRASIL, P., RODRIGUES, R.R., ROVARIS, D.B., FERNANDES, S.B., FERNANDES, C., SANTOS, J.H.A., ABDALLA, L.F., COSTA-FILHO, R., SILVA, M., SOUZA, V., COSTA, Á.A., MEJÍA, M., BRANDÃO, M.J., GONÇALVES, L.F., SILVA, G.A., DE JESUS, M.S., PESSOA, K., CORADO, A.L.G., DUARTE, D.C.G., MACHADO, A.B., ZUKERAM, K.A., VALENTE, N., LOPES, R.S., PEREIRA, E.C., APPOLINARIO, L.R., ROCHA, A.S., TORT, L.F.L., SEKIZUKA, T., ITOKAWA, K., HASHINO, M., KURODA, M., DEZORDI, F.Z., WALLAU, G.L., DELATORRE, E., GRÄF, T., SIQUEIRA, M.M., BELLO, G. and RESENDE, P.C., 2023. SARS-CoV-2 intra-host diversity, antibody response, and disease severity after reinfection by the variant of concern Gamma in Brazil. *Scientific Reports*, vol. 13, no. 1, pp. 7306. http://doi.org/10.1038/s41598-023-33443-1. PMid:37147348.

NELSON, G., BUZKO, O., SPILMAN, P., NIAZI, K., RABIZADEH, S., & SOON-SHIONG, P., 2021. Molecular dynamic simulation reveals E484K mutation enhances spike RBD-ACE2 affinity and the combination of E484K, K417N and N501Y mutations (501Y.V2 variant) induces conformational change greater than N501Y mutant alone, potentially resulting in an escape mutant. *BioRxiv*, In press. https://doi.org/10.1101/2021.01.13.426558.

PENETRA, S.L.S., SANTOS, H.F.P., RESENDE, P.C., BASTOS, L.S., DA SILVA, M.F.B., PINA-COSTA, A., LOPES, R.S., SABOIA-VAHIA, L., DE OLIVEIRA, A.C.A., PEREIRA, E.C., FILHO, F.M., WAKIMOTO, M.D., CALVET, G.A., FULLER, T.L., WHITWORTH, J., SMITH, C., NIELSEN-SAINES, K., CARVALHO, M.S., ESPÍNDOLA, O.M., GUARALDO, L., SIQUEIRA, M.M. and BRASIL, P., 2023. SARS-CoV-2 reinfection cases in a household-based prospective cohort in Rio de Janeiro. *The Journal of Infectious Diseases*, vol. 228, no. 12, pp. 1680-1689. http://doi.org/10.1093/infdis/jiad336. PMid:37571849.

POSADA, D., 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, vol. 25, no. 7, pp. 1253-1256. http://doi.org/10.1093/molbev/msn083. PMid:18397919.

RAMBAUT, A., HOLMES, E.C., O'TOOLE, Á., HILL, V., MCCRONE, J.T., RUIS, C., DU PLESSIS, L. and PYBUS, O.G., 2020a. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology*, vol. 5, no. 11, pp. 1403-1407. http://doi.org/10.1038/s41564-020-0770-5. PMid:32669681.

RAMBAUT, A., LOMAN, N., PYBUS, O., BARCLAY, W., BARRETT, J., CARABELLI, A., CONNOR, T., PEACOCK, T., ROBERTSON, D.L., VOLZ, E., 2020b [viewed 27 November 2023]. *Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations* [online]. Edinburgh: Virological. Available from: https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563

SANTOS, L.A., GÓIS FILHO, P.G., SILVA, A.M.F., SANTOS, J.V.G., SANTOS, D.S., AQUINO, M.M., JESUS, R.M., ALMEIDA, M.L.D., SILVA, J.S., ALTMANN, D.M., BOYTON, R.J., SANTOS, C.A., SANTOS, C.N.O., ALVES, J.C., SANTOS, I.L., MAGALHÃES, L.S., BELITARDO, E.M.M.A., ROCHA, D.J.P.G., ALMEIDA, J.P.P., PACHECO, L.G.C., AGUIAR, E.R.G.R., CAMPOS, G.S., SARDI, S.I., CARVALHO, R.H., JESUS, A.R., REZENDE, K.F. and ALMEIDA, R.P., 2021. Recurrent COVID-19 including evidence of reinfection and enhanced severity in thirty Brazilian healthcare workers. *The Journal of Infection*, vol. 82, no. 3, pp. 399-406. http://doi.org/10.1016/j.jinf.2021.01.020.

SPIRA, B., 2022. The impact of the highly virulent SARS-CoV-2 Gamma variant on young adults in the state of São Paulo: was it inevitable? *Cureus*, vol. 14, no. 7, e26486. http://doi.org/10.7759/cureus.26486. PMid:35919213.

THAKUR, V., BHOLA, S., THAKUR, P., PATEL, S.K.S., KULSHRESTHA, S., RATHO, R.K. and KUMAR, P., 2022. Waves and variants of SARS-CoV-2: understanding the causes and effect of the COVID-19 catastrophe. *Infection*, vol. 50, no. 2, pp. 309-325. http://doi.org/10.1007/s15010-021-01734-2. PMid:34914036.

THE LANCET, 2021. Genomic sequencing in pandemics. *The Lancet*, vol. 397, no. 10273, pp. 445. https://doi.org/10.1016/S0140-6736(21)00257-9

THERMO FISHER SCIENTIFIC, 2023 [viewed 25 February 2023]. *Ion AmpliSeq TM SARS-CoV-2 Insight Research Assay Product description* [online]. Available from: www.thermofisher.com/us/en/home/global/terms-and-conditions.html

VOLZ, E., MISHRA, S., CHAND, M., BARRETT, J.C., JOHNSON, R., GEIDELBERG, L., HINSLEY, W.R., LAYDON, D.J., DABRERA, G., O'TOOLE, Á., AMATO, R., RAGONNET-CRONIN, M., HARRISON, I., JACKSON, B., ARIANI, C.V., BOYD, O., LOMAN, N.J., MCCRONE, J.T., GONÇALVES, S., JORGENSEN, D., MYERS, R., HILL, V., JACKSON, D.K., GAYTHORPE, K., GROVES, N., SILLITOE, J., KWIATKOWSKI, D.P., FLAXMAN, S., RATMANN, O., BHATT, S., HOPKINS, S., GANDY, A., RAMBAUT, A. and FERGUSON, N.M., 2021. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature*, vol. 593, no. 7858, pp. 266-269. http://doi.org/10.1038/s41586-021-03470-x. PMid:33767447.

WORLD HEALTH ORGANIZATION – WHO., 2024 [viewed 12 January 2024]. *WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data* [online]. Available from: https://covid19.who.int/

YAHAV, D., YELIN, D., ECKERLE, I., EBERHARDT, C.S., WANG, J., CAO, B. and KAISER, L., 2021. Definitions for coronavirus disease 2019 reinfection, relapse and PCR re-positivity. *Clinical Microbiology and Infection*, vol. 27, no. 3, pp. 315-318. http://doi.org/10.1016/j.cmi.2020.11.028. PMid:33285276.

ZHAO, Y., HUANG, J., ZHANG, L., CHEN, S., GAO, J. and JIAO, H., 2022. The global transmission of new coronavirus variants. *Environmental Research*, vol. 206, pp. 112240. http://doi.org/10.1016/j.envres.2021.112240. PMid:34688639.

ZHOU, B., THAO, T.T.N., HOFFMANN, D., TADDEO, A., EBERT, N., LABROUSSAA, F., POHLMANN, A., KING, J., STEINER, S., KELLY, J.N., PORTMANN, J., HALWE, N.J., ULRICH, L., TRÜEB, B.S., FAN, X., HOFFMANN, B., WANG, L., THOMANN, L., LIN, X., STALDER, H., POZZI, B., DE BROT, S., JIANG, N., CUI, D., HOSSAIN, J., WILSON, M.M., KELLER, M.W., STARK, T.J., BARNES, J.R., DIJKMAN, R., JORES, J., BENARAFA, C., WENTWORTH, D.E., THIEL, V. and BEER, M., 2021. SARS-CoV-2 spike D614G change enhances replication and transmission. *Nature*, vol. 592, no. 7852, pp. 122-127. http://doi.org/10.1038/s41586-021-03361-1. PMid:33636719.

ZIMERMAN, R.A., FERRAREZE, P.A.G., CADEGIANI, F.A., WAMBIER, C.G., FONSECA, D., SOUZA, A.R., GOREN, A., ROTTA, L.N., REN, Z. and THOMPSON, C.E., 2022. Comparative Genomics and Characterization of SARS-CoV-2 P.1 (Gamma) variant of concern from Amazonas, Brazil. *Frontiers in Medicine*, vol. 9, pp. 806611. http://doi.org/10.3389/fmed.2022.806611. PMid:35242782.

Supplementary Material

Supplementary material accompanies this paper.

Supplementary Material 1 - Sample IDs.

The raw data supporting the conclusions of this article are available on the GISAID platform (https://gisaid.org/). Sample IDs can be found in the supplementary material.

This material is available as part of the online article from https://doi.org/10.1590/1519-6984.292021.