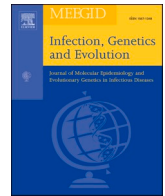




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Research paper

SARS-CoV-2 introduction and lineage dynamics across three epidemic peaks in Southern Brazil: massive spread of P.1

Ana Paula Mutterle Varela^{a,1}, Janira Prichula^{a,1}, Fabiana Quoos Mayer^b,
Richard Steiner Salvato^c, Fernando Hayashi Sant'Anna^d, Tatiana Schäffer Gregianini^c,
Letícia Garay Martins^e, Adriana Seixas^a, Ana Beatriz Gorini da Veiga^{a,*}

^a Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Rua Sarmento Leite, 245, 90050-170 Porto Alegre, Rio Grande do Sul, Brazil

^b Centro de Pesquisa Em Saúde Animal, Instituto de Pesquisas Veterinárias Desidério Finamor, Departamento de Diagnóstico e Pesquisa Agropecuária, Secretaria da Agricultura, Pecuária e Desenvolvimento Rural, Estrada Do Conde, 6000, Sans Souci, 92990-000 Eldorado do Sul, Rio Grande do Sul, Brazil

^c Laboratório Central de Saúde Pública, Centro Estadual de Vigilância em Saúde da Secretaria de Saúde do Estado do Rio Grande do Sul (LACEN/CEVS/SES-RS). Av. Ipiranga, 5400, 90610-000 Porto Alegre, Rio Grande do Sul, Brazil

^d Hospital Moinhos de Vento, PROADI – SUS, Rua Ramiro Barcelos, 910 – Moinhos de Vento, Porto Alegre, Rio Grande do Sul, Brazil

^e Centro Estadual de Vigilância em Saúde da Secretaria de Saúde do Estado do Rio Grande do Sul – CEVS/SES-RS. Av. Ipiranga, 5400, 90610-000 Porto Alegre, Rio Grande do Sul, Brazil



ARTICLE INFO

Keywords:

Severe acute respiratory syndrome coronavirus 2
Genomic diversity
P.1 lineage
Epidemic peaks
Brazil

ABSTRACT

Genomic surveillance of SARS-CoV-2 is paramount for understanding viral dynamics, contributing to disease control. This study analyzed SARS-CoV-2 genomic diversity in Rio Grande do Sul (RS), Brazil, including the first reported case in each Regional Health Coordination and cases from three epidemic peaks. Ninety SARS-CoV-2 genomes from RS were sequenced and analyzed through comparison with SARS-CoV-2 datasets available in GISAID for phylogenetic inference and mutation analysis. Among the first reported cases, we found the following lineages: B.1 (33.3%), B.1.1.28 (26.7%), B.1.1 (13.3%), B.1.1.33 (10.0%), and A (6.7%), evidencing SARS-CoV-2 introduction by both international origin and community-driven transmission. We found predominance of B.1.1.33 (50.0%) and B.1.1.28 (35.0%) during the first epidemic peak (July–August 2020), emergence of P.2 (55.6%) in the second peak (November–December 2020), and massive spread of P.1 and related sequences (78.4%), such as P.1-like-II, P.1.1 and P.1.2 in the third peak (February–April, 2021). Eighteen novel mutation combinations were found among P.1 genomes, and 22 different spike mutations and/or deletions among P.1 and related sequences. This study shows the dispersion of SARS-CoV-2 lineages in Southern Brazil and describes SARS-CoV-2 diversity during three epidemic peaks, highlighting the spread of P.1 and the high genetic diversity of currently circulating lineages. Genomic monitoring of SARS-CoV-2 is essential to guide health authorities' decisions to control COVID-19 in Brazil.

1. Introduction

The new coronavirus SARS-CoV-2 emerged at the end of 2019 in the province of Wuhan, China, and rapidly spread to other countries, infecting millions of people worldwide (Platto et al., 2020; Zhu et al., 2020). In Brazil, as of 22 September 2021 approximately 21.2 million cumulative cases and 591,000 deaths have been reported, ranking third country in the cumulative number of cases and second in the number of deaths (Johns Hopkins Coronavirus Resource Center, 2021). The first

confirmed case in Brazil was a man returning from Italy to São Paulo on 26 February 2020; following that, different viral strains were introduced in the country by individuals returning from international travel (Ministério da Saúde, 2021). Notably, with high viral transmission, variants of concern (VOC), variants of interest (VOI) and variants under investigation (VUI) emerged in the Brazilian population, such as P.1, P.2 and VUI-NP13L, respectively, all derived from B.1.1.28 (Francisco et al., 2021; Naveca et al., 2021; Faria et al., 2021; Voloch et al., 2021).

Rio Grande do Sul (RS), the southernmost Brazilian state, reports

* Corresponding author at: Departamento de Ciências Básicas da Saúde, Universidade Federal de Ciências da Saúde de Porto Alegre, Rua Sarmento Leite, 245, CEP 90050-170 Porto Alegre, RS, Brazil.

E-mail address: anabgv@ufcspa.edu.br (A.B.G. Veiga).

¹ These authors have contributed equally to this work and share first authorship.

<https://doi.org/10.1016/j.meegid.2021.105144>

Received 26 September 2021; Received in revised form 9 November 2021; Accepted 14 November 2021

Available online 17 November 2021

1567-1348/© 2021 The Authors.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

high numbers of respiratory viral infections annually (Gorini da Veiga et al., 2012). RS is a hub for connecting flights, with four international airports and 56 regional airports, seven of which have regular routes; moreover, RS borders Uruguay and Argentina, is close to Paraguay, and is a popular tourist destination in Brazil (Instituto Brasileiro de Geografia e Estatística, 2021). Hence, RS is one of the Brazilian states with the highest number of reported SARS-CoV-2 variants (Secretaria da Saúde do Estado do Rio Grande do Sul, 2021).

In RS, there are 19 Regional Health Coordinations (RHC) responsible for epidemiological surveillance, whose data show three COVID-19 epidemic peaks in the state until June 2021 (Secretaria da Saúde do Estado do Rio Grande do Sul, 2021). The first peak occurred in August 2020, when the cumulative number of cases per day was around 3000; the second peak was in November–December 2020, with approximately 6000 cases reported daily; the third and more critical peak started in February and lasted until April 2021, in which more than 13,000 cases were reported in a single day (Fig. 1). Notably, in March 2021 RS became the epicenter of COVID-19 in Brazil, with great health and economic burden, and exhaustion of the health system (Secretaria da Saúde do Estado do Rio Grande do Sul, 2021).

To better understand the spread of SARS-CoV-2 lineages in RS, we performed genomic analysis of SARS-CoV-2 from the first cases reported in 15 RHCs, in addition to samples from three epidemic peaks. Our findings help to understand SARS-CoV-2 transmission and diversity in Southern Brazil, contributing to the Brazilian Public Healthcare System.

2. Materials and methods

2.1. Study design

Ninety upper- and lower-respiratory tract secretion samples from patients with SARS-CoV-2 infection confirmed by RT-qPCR were obtained from the State Central Laboratory of Rio Grande do Sul (LACEN-RS) (Supplementary Data 1). To investigate spread and diversity of SARS-CoV-2 lineages across RS we selected the first positive case from 15 RHCs of RS ($n = 15$) (Supplementary Table 1), and samples ($n = 75$) from the four cities with the highest cumulative incidences (Santo Angelo, Passo Fundo, Caxias do Sul, and Porto Alegre), comprising three epidemic peaks: July–August 2020 (first peak; $n = 20$), November–December 2020 (second peak; $n = 18$), and February–April 2021 (third peak; $n = 37$) (Fig. 1). This study was approved by the Ethics Committee of Universidade Federal de Ciências da Saúde de Porto Alegre (Protocol n. 3.978.647, CAAE 30714520.0.0000.5345).

2.2. RNA extraction and sequencing

Viral RNA was extracted with MagMAX™ Viral/Pathogen Nucleic Acid Isolation kit in a KingFisher™ Flex Purification System (ThermoFisher Scientific). SARS-CoV-2 RT-qPCR was performed on a 7500 Real-Time PCR System (Applied Biosystems) using Seegene-Allplex 2019-nCoV Assay, following the manufacturer's instructions.

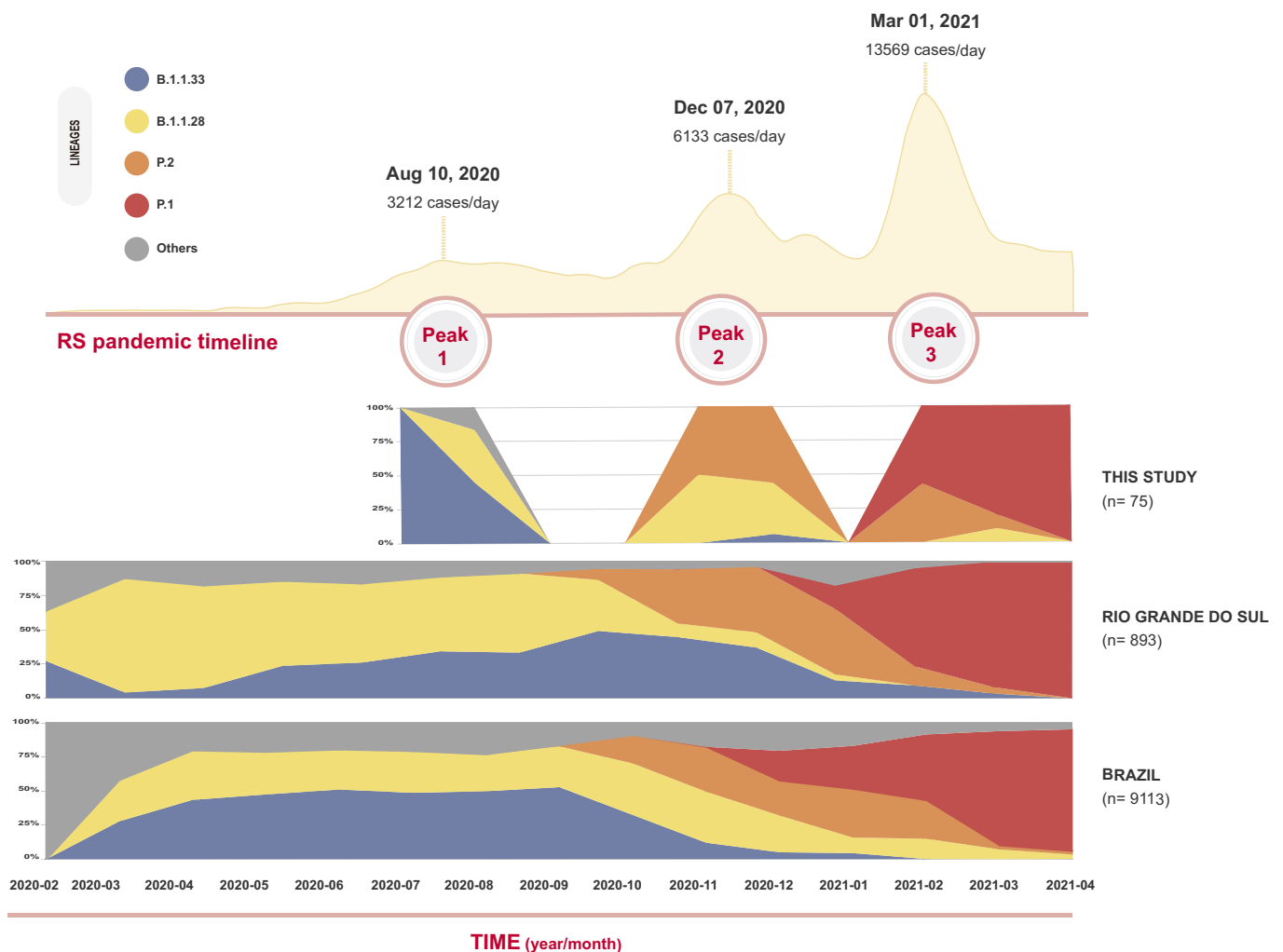


Fig. 1. Profile of SARS-CoV-2 lineage distribution from March 2020 to April 2021. Cumulative number of COVID-19 confirmed cases in RS state during three peaks (July–August 2020; November–December 2020; and February–April 2021), and density distribution of the main lineages identified in this study.

Libraries were constructed using QIASEQ SARS-CoV-2 Primer Panel and QIAseq FX DNA Library UDI-A kit (Qiagen) following manufacturer's instructions, with previously described annealing temperature (Sant'Anna et al., 2021). Libraries were quantified using Qubit™ dsDNA HS Assay kit and normalized to equimolar concentrations. Sequencing was performed with MiSeq Reagent Kit v3 600 cycles in a MiSeq instrument (Illumina, USA).

2.3. Bioinformatic analysis

Raw fastq files were trimmed to remove adapter sequences and low-quality reads using trimmomatic (Bolger et al., 2014) (Parameters: leading:3 trailing:3 slidingwindow:4:25 minlen:36 illuminaclip:TruSeq3-PE.fa:2:30:10). Reads were mapped against SARS-CoV-2 reference genome (GenBank Accession NC_045512.2) using the bwa-mem algorithm (v0.7.17) (Li and Durbin, 2009) under default parameters. The quality of the resulting mapped BAM files was checked using Qualimap (García-Alcalde et al., 2012). The software iVar (v1.2.2) was used to trim QIASEQ SARS-CoV-2 Primer Panel sequences and along with Samtools (v1.6) mpileup function (Li et al., 2009) to variant calling (parameters: -A -d 20,000 -Q 0) and consensus sequence generation (parameters: -q 20 -t 0). Consensus sequences were submitted to Next-clade (<https://clades.nextstrain.org>) for quality inspection and to determine the mutational profile. Sequences were submitted to Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin) v2.4.2 (<https://github.com/cov-lineages/pangolin>) for lineage assignment. Mutations related to lineage P.1 were also manually inspected for assessment of lineages P.1.1, P.1.2 and P.1-like (Gräf et al., 2021; Almeida et al., 2021). The mutation profiles of P.1, P.1.1, P.1.2 and P.1-likeII genomes were subjected to covSPECTRUM to evaluate new mutation combinations. The heatmap with dendrogram was constructed for mutation profile visualization using Heatmapper (Babicki et al., 2016).

2.4. Temporal lineage dynamic and phylogenomic analysis

To evaluate the dynamic of SARS-CoV-2 lineages in RS across the epidemic peaks, the relative density profile was accessed using sequenced SARS-CoV-2 genomes combined with two different SARS-CoV-2 datasets available in the EpiCoV database in GISAID as of 18 May 2021: 893 sequences from RS, and 9113 sequences from other Brazilian regions. In addition, two phylogenomic analysis were performed, one concatenating our dataset of 90 genome sequences with the Nextstrain' South America dataset (2961 sequences) and other concatenating only sequences classified as P.1 and P.1-like from Brazil available in GISAID (between December 2019 and April 2021). Analyses were carried out using Nextstrain ncov workflow (<https://github.com/>) (default parameters), which includes alignment, phylogenetic reconstruction, geographic and ancestral trait reconstruction, and inference of transmission events. Nextstrain uses by default the software IQ-Tree and the substitution model GTR. For the P.1 phylogenomic reconstruction, we adopted a subsampling scheme focused on the RS (group by division, year, and month, maintaining 20 sequences per group).

3. Results

3.1. Clinical parameters

Of the 90 samples analyzed in this study, 47 were from females (median age 39 years [1–105 years]), and 43 from males (median age 52 years [6 months–91 years]). Based on available data, 42.2% of the patients (35/83) were hospitalized, and 23.8% (15/63) died. Cough, fever, and dyspnea were the main symptoms, whereas cardiopathy and diabetes were the main reported comorbidities (Supplementary Table 2).

3.2. Sequencing data and SARS-CoV-2 lineages diversity

The number of paired-end reads per sequenced genome varied from 94,300 to 697,700, with mean depth coverage ranging from 43× to 558× (Supplementary Data 1), and presenting at least 99% coverage breadth to the Wuhan-Hu-1 reference genome (NC_045512.2).

According to Pangolin and to manual verification based on previously described mutational profile (Naveca et al., 2021; Faria et al., 2021; Gräf et al., 2021; Almeida et al., 2021), our SARS-CoV-2 genomes were assigned to 11 lineages: P.1 ($n = 20$; 22.2%), B.1.1.28 ($n = 19$; 21.1%), P.2 ($n = 17$; 18.9%), B.1.1.33 ($n = 15$; 16.7%), P.1.2 ($n = 7$; 7.8%), B.1 ($n = 5$; 5.6%), B.1.1 ($n = 3$; 3.3%), B.1.91 ($n = 1$; 1.1%), A ($n = 1$; 1.1%), P.1.1 ($n = 1$; 1.1%) and P.1-likeII ($n = 1$; 1.1%) (Supplementary Data 1). Density profile analysis shows that the main lineages identified in this study are in accordance with the lineage distribution observed for RS and Brazil (Fig. 1).

Among the first reported cases of 15 RHCs, we identified lineages A ($n = 1$; 6.7%), B.1 ($n = 5$; 33.3%), B.1.1 ($n = 2$; 13.3%), B.1.1.28 ($n = 4$; 26.7%), B.1.1.33 ($n = 3$; 20.0%) (Fig. 2A and B). Notably, lineages A and B.1 were detected in samples from March 2020, whereas the others were identified between March and May 2020 (Supplementary Data 1).

Following SARS-CoV-2 introduction in the state, the first peak was marked by lineages B.1 ($n = 1$; 5.0%), B.1.1.28 ($n = 7$; 35.0%), B.1.1.33 ($n = 11$; 55.0%), and B.1.91 ($n = 1$; 5.0%). In the second peak, P.2 was the most frequent lineage ($n = 10$; 55.6%), followed by B.1.1.28 ($n = 7$; 38.9%) and B.1.1.33 ($n = 1$; 5.6%). In the third peak, P.1 was the most frequent lineage ($n = 20$; 54.1%), followed by P.1.2 ($n = 7$; 18.9%), P.2 ($n = 7$; 18.9%), P.1.1 ($n = 1$; 2.7%), B.1.1.28 ($n = 1$; 2.7%) and the recently described P.1-likeII ($n = 1$; 2.7%) (Figs. 1 and 2B).

3.3. Phylogenomic analysis and mutations

Phylogenetic reconstruction enriched to the South American dataset revealed that the RS genomes grouped according to the Pangolin classification (Fig. 3). Accordingly, phylogenetic analysis showed that the lineage A sequence from this study grouped with other lineage A sequences from Peru in a major clade containing genomes of lineage A.5 from Bolivia, Uruguay, Chile, and Peru. In B.1.1.28 clade, we identified three genomes (85,301, 99,718 and 22,405) grouped with VUI-NP13L, confirmed by the presence of previously reported amino acid change combination (Francisco et al., 2021; Sant'Anna et al., 2021). Regarding lineage P.2, seven genomes clustered with P.2 sequences from Uruguay, Australia, and England.

Based on phylogenetic tree (Fig. 4), 29 sequences classified as P.1 and P.1-related were divided in two main groups: one containing sequences of P.1, P.1.1 and P.1.2 ($n = 28$), and other group with the P.1-likeII sequence ($n = 1$). The P.1.1 sequence, recovered in this study, clustered with two sequences from São Paulo, while the P.1.2 sequences grouped in a subcluster containing two other genomes recovered from São Paulo and RS (Fig. 4).

Comparing the genomes sequenced in this study with the Wuhan-Hu-1 reference genome (NC_045512.2), a total of 416 nucleotide substitutions, 8 deletions and 1 insertion were found. These mutations comprised 247 amino acid (aa) substitutions and 7 aa deletions: ORF1a (89), ORF1b (47), S (36), ORF3a (21), M (4), ORF6 (6), ORF7a (6), ORF7b (2), ORF8 (13), ORF9b (5), and N (25) (Supplementary Data 1).

Regarding P.1 and P.1-related genomes ($n = 29$), 87 non-synonymous mutations were detected (82 substitutions; 5 deletions), 20 of which are present in all sequences (Supplementary Table 3). These mutations were found in ORF1a (32), ORF1b (10), S (22), ORF3a (6), M (1), ORF7a (4), ORF7b (1), ORF8 (3), ORF9b (3) and N (5) (Fig. 5A). In the Spike gene, the ten lineage-defining mutations (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, H655Y and T1027I) were confirmed, and 12 additional mutations/deletions were found: L5F, V1176F, N188S, L189-, R190-, T573I, D614G, E661D, P812S, A845S, T859I, and V1264L. Based on covSPECTRUM analysis, 17 new mutation

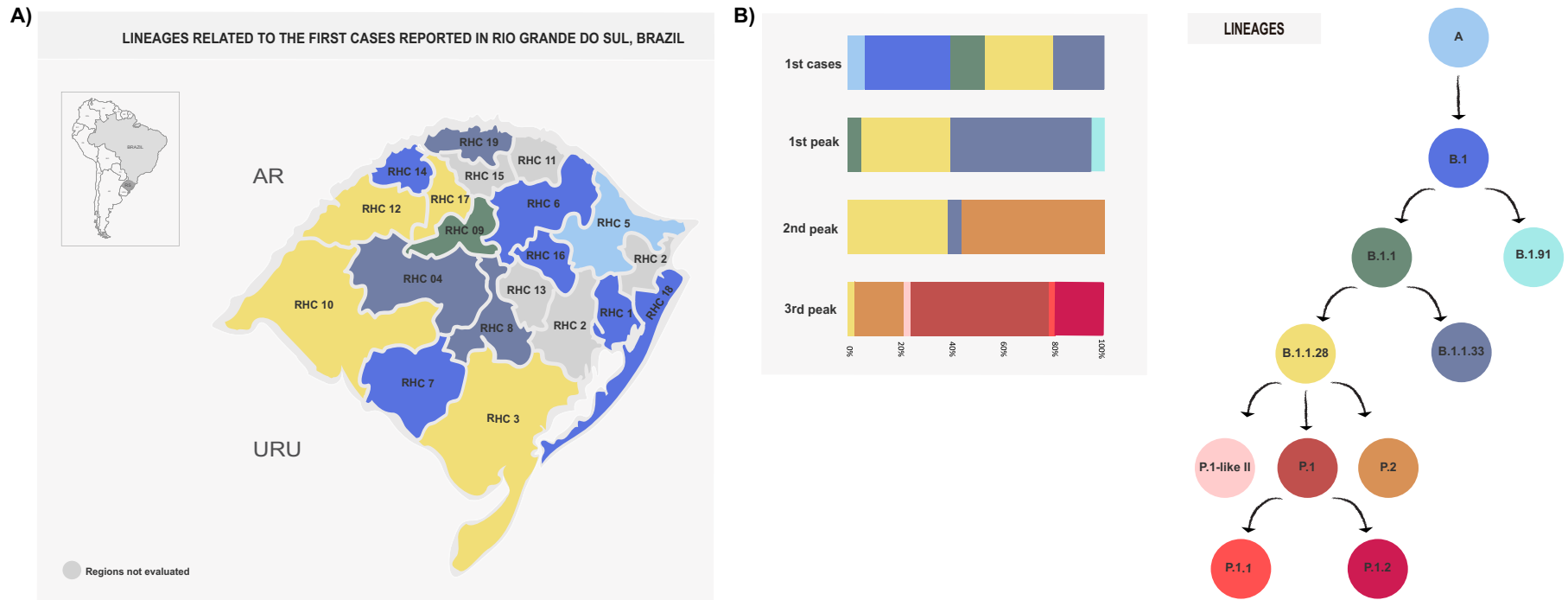


Fig. 2. SARS-CoV-2 lineages diversity in Rio Grande do Sul (RS), Brazil. A. The inside box shows South America, with Brazil in light grey and the state of RS in dark grey. The larger map shows the 19 Regional Health Coordinations of RS, among which 15 are colored according to the SARS-CoV-2 lineage identified in this study. B. Profile of SARS-CoV-2 lineages identified in the first reported cases and in the three epidemic peaks in RS. Moreover, a schematic representation based on lineages' phylogenetic relationship was used as legend.

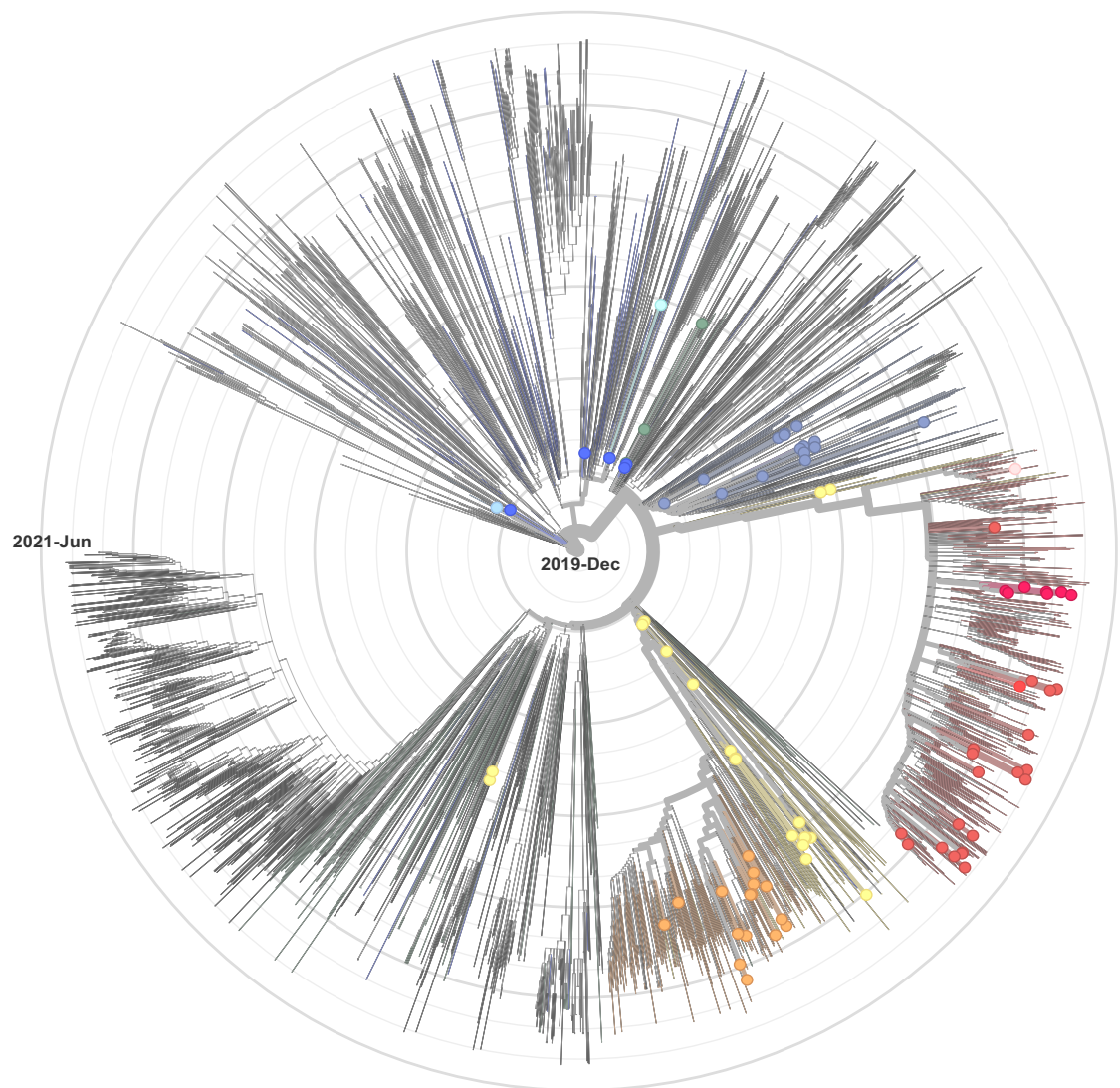


Fig. 3. Phylogenomic tree of SARS-CoV-2 genomes. Sequences obtained in this study were combined with a dataset of 2961 South American genomes. Our SARS-CoV-2 sequences are colored according to lineage classification.

combination profiles were detected in 18 P.1-related genomes (Fig. 5B).

Three P.1 genomes, recovered from Santo Angelo city in the third peak, showed the mutations Δ L189, Δ R190 and N188S in the Spike, grouping in the heatmap dendrogram (Fig. 5C). In the phylogenetic analysis, these sequences clustered with other SARS-CoV-2 genomes from Brazil (Bahia, Roraima, Goias and RS) and from Colombia and French Guiana (Figs. 3 and 4).

4. Discussion

The COVID-19 pandemic has deeply impacted the health system and the economy of Brazil, overcrowding hospitals and overburdening health professionals. Genomic epidemiology has been paramount in understanding SARS-CoV-2 dispersion and to track the evolutive dynamics of viral transmission. In view of this, we investigated the viral dynamics of SARS-CoV-2 in RS from the first cases reported and across the three main peaks of infection that occurred between July 2020 and April 2021 in the region.

The first individual with SARS-CoV-2 infection confirmed in RS had returned from Italy in early March 2020 and our analysis identified it as of lineage B.1. Other three introductory cases were also B.1. Traveling records obtained from these cases showed that two patients returned

from Europe, whereas the third had been in Rio de Janeiro. Notably, B.1 was the lineage related to the early epidemic outbreak in northern Italy (Lai et al., 2020; Zehender et al., 2020). Hence our findings confirm this route of international introduction of SARS-CoV-2 in RS. In addition, B.1.1 lineage, which evolved from B.1 and circulated in Europe in the beginning of the pandemic, was also found among the first cases reported in RS.

The lineage A was also found among the first reported cases, in a resident of Serafina Correa city who had been in Paraguay; in the phylogenetic analysis, this sample clustered with sequences from Peru. According to Pangolin data, lineage A is considered to be one of the two major lineages of the root of the pandemic, originating in China and spreading to Asia, Europe, Oceania, and North America. In South America, lineage A was detected early in the pandemic at low frequencies, with only a few genome sequences reported in Brazil, all from March 2020.

Besides these variants, Brazilian lineages B.1.1.28 and B.1.1.33, which emerged in São Paulo in late February 2020 (Candido et al., 2020; Resende et al., 2021a), were also detected among the first cases reported in RS. The dispersion pattern and demographical dynamics of these lineages showed that they were introduced by multiple events, and their spread occurred throughout Brazil by community transmission in early

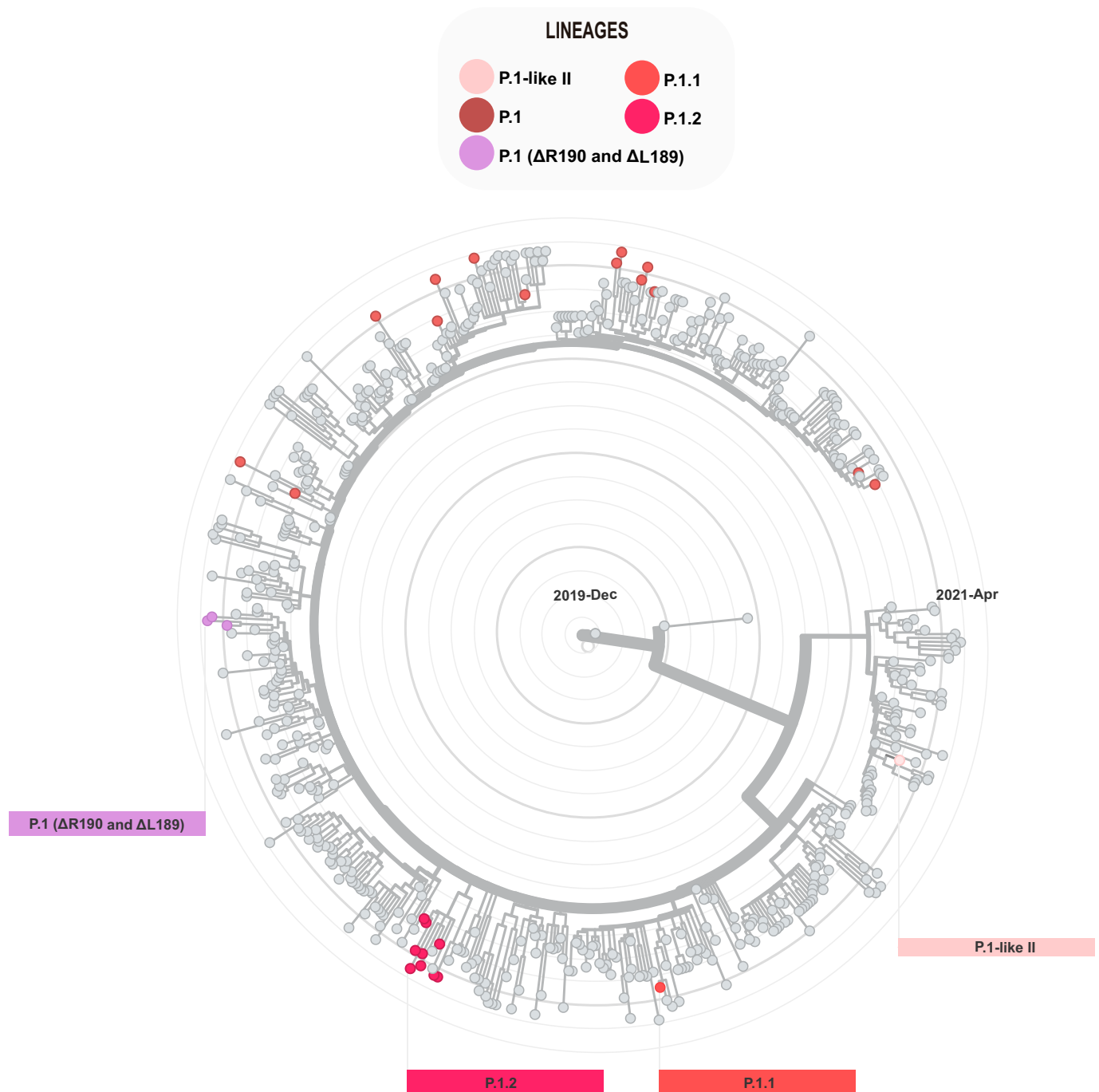


Fig. 4. Phylogenomic reconstruction of P.1 and P.1-related lineages. Our SARS-CoV-2 genomes were combined with a dataset of Brazilian P.1, P.1.1, P.1.2 and P.1-like genomes. Sequences recovered in this study are colored according to lineage classification.

March, and also to other countries (Candido et al., 2020; Franceschi et al., 2021; Mir et al., 2021; Resende et al., 2021a).

Our genomic analyses of the first SARS-CoV-2-confirmed cases in RS suggests that viral introduction in the state was related to both international origin and community-driven transmission. In March 2020, international travel restrictions were adopted to control viral spread, influencing SARS-CoV-2 transmission pattern. Accordingly, previous studies show decreases in imported cases after travel restrictions in Brazil, which, in turn, contribute to increased circulation of local lineages (Candido et al., 2020; Resende et al., 2021a).

Additionally, our study reveals the temporal frequency and divergence of SARS-CoV-2 lineages in RS over time. As observed in previous Brazilian studies, B.1.1.28 and B.1.1.33 predominated during the first

epidemic peak (July and August 2020) (Candido et al., 2020; Francisco et al., 2021; Sant'Anna et al., 2021; Voloch et al., 2021). Selection of these lineages over others might be explained by mutations associated with higher viral fitness and severe disease (Nagy et al., 2021). Moreover, the V1176F mutation in B.1.1.28 is predicted to increase the flexibility of the stalk domain in the Spike protein trimmer, facilitating its binding to the ACE2 receptor (Farkas et al., 2020).

B.1.1.28 and B.1.1.33 continued to circulate in the second peak; however, an emergence of P.2 was observed. This lineage descends from B.1.1.28, carrying the lineage-defining mutations ORF1ab:L3468V, ORF1ab:synC11824U, N:A119S, and S:E484K (Voloch et al., 2021). According to our previous study, P.2 has been massively circulating in RS since October 2020 (Sant'Anna et al., 2021). Additionally, in the

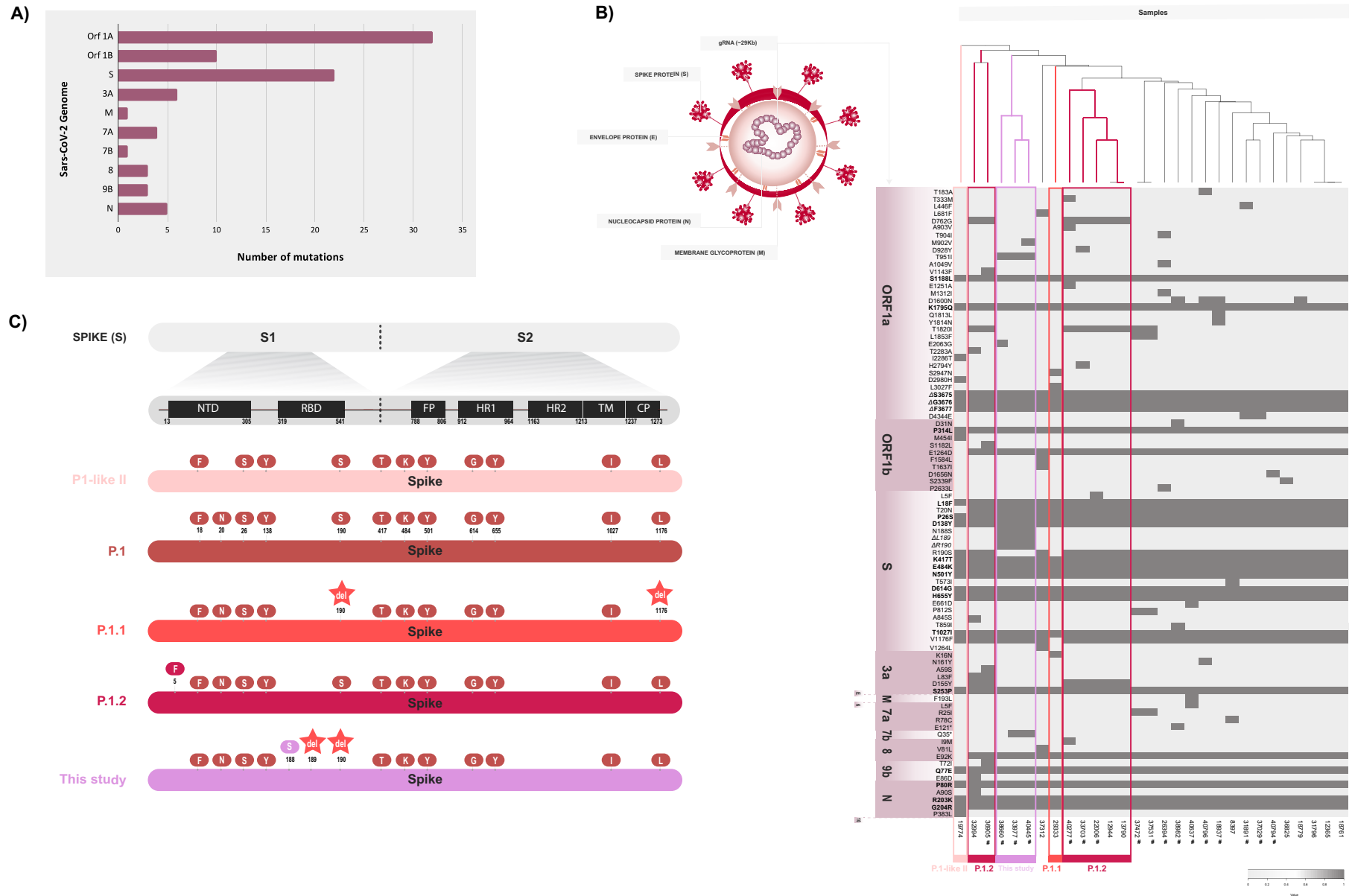


Fig. 5. SARS-CoV-2 mutation patterns over time present in P1 and P1-related genomes of this study compared to the reference Wuhan-Hu-1 sequence. **A.** Total of mutations across encoding regions of 29 SARS-CoV-2 genomes of P.1, P.1.1, P.1.2, and P.1-likeII. **B.** Heatmap displaying the non-synonymous mutations present in P.1 and P.1-related sequences. These mutations are plotted according to the respective genome position and are summarized as presence (dark grey) or absence (light grey). A dendrogram based on mutation profile and a schematic representation of SARS-CoV-2 virion structure are also exhibited. **C.** Scheme of mutations on spike proteins highlighting the differences in the protein profile among P.1 and P.1-related lineages.

present study three genomes (two in the second peak and one in third peak) classified as B.1.1.28 were shown to be VUI-NP13L, a potential new lineage (Francisco et al., 2021; Sant'Anna et al., 2021). VUI-NP13L was first reported in RS in August 2020, disseminating afterwards (Sant'Anna et al., 2021), and our results reinforce that this lineage continued to circulate at low frequency in RS.

The third peak was marked by P.1 and P.2 as the predominant lineages. Both descend from B.1.1.28 have distinct origins, and share the S: E484K mutation (Harvey et al., 2021). P.1 was first identified in four travelers arriving in Japan from Amazonas, Brazil (on 2 January 2021) (Fujino et al., 2021), and was later confirmed in several Amazonas samples (in November 2020) (Faria et al., 2021). In our study, P.1 and P.1-related lineages were the most frequent in RS since the first week of February 2021, becoming the dominant lineages in the third peak; P.1 was first identified in late January 2021 in Gramado, a touristic town that annually receives 6.5 million visitors (Salvato et al., 2021). The P.1 rapid dispersion might be related to its higher transmission rate (Coutinho et al., 2021), higher viral loads (Naveca et al., 2021), greater affinity to ACE2 receptor (Dejnirattisai et al., 2021), and the ability to resist neutralizing antibodies, either from natural infections or vaccine-induced (Dejnirattisai et al., 2021; Harvey et al., 2021; Hoffmann et al., 2021). Moreover, P.1 spread may also be influenced by relaxed social distancing measures at holidays and vacation, especially during the summer period in the state.

P.1 lineage became a VOC based on lineage-defining mutations in proteins S (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, H655Y and T1027I), NSP3 (S370L and K977Q), NSP13 (E341D), NS8 (E92K) and N (P80R) (Gräf et al., 2021; Naveca et al., 2021). The high number of lineage-defining mutations for P.1 is considered a result of sequential infection steps during B.1.1.28 evolution in Amazonas (Gräf et al., 2021). Our results show the expressive spread of P.1 across Brazilian states and the high genetic diversity of P.1-related lineages in RS since February 2021. Besides lineage-defining mutations, 87 additional mutations were found among our P.1 genomes. Moreover, 17 new mutation combinations were observed in 18 genomes, suggesting that independent events are related to the observed variants. Accordingly, the P.1 genomes obtained here are present in several clades in the phylogenomic tree; additionally, this rapid diversification was evidenced by P.1.1 and P.1.2 detection.

Among our P.1 sequences, three genomes displayed two deletions in the S gene (Δ L189 and Δ R190) in the N-terminal domain (NTD), which were recently found in four other genomes recovered in Brazil (Resende et al., 2021b). Deletions in NTD have been reported during prolonged infection of immunocompromised patients (McCarthy et al., 2021; Ribes et al., 2021) and subsequent transmission (Hensley et al., 2021). McCarthy et al. (2021) observed deletions in the S gene, of which >97% maintained the open reading frame, and 90% occurred in four sites, named "recurrent deletion regions 1–4" (RDRs). Considering these RDRs, the Δ L189 and Δ R190 deletions are located between RDR2 (position 139–146) and RDR3 (position 210–212). Although both deletions were not reported to affect the NTD antigenic-supersite, they might lead to conformational changes in exterior loops, affecting antibody binding outside the antigenic-supersite (Resende et al., 2021b). Deletions in NTD have been associated with resistance to antibody neutralization, suggesting an improvement in virus fitness by evading the host's immune response, an evolutive process due to immune pressure (McCarthy et al., 2021; Resende et al., 2021b; Ribes et al., 2021).

Further *in vitro* and *in vivo* studies are needed to address the role of these mutations in viral fitness. This is important for predicting and understanding viral infectivity, pathogenicity, and resistance to vaccines. In this sense, continuous genomic monitoring of SARS-CoV-2 is essential to identify new amino acid changes and evaluate their impact on lineages spreading in Brazil.

5. Conclusion

This study allowed the analysis of the first SARS-CoV-2 lineages and of the main epidemic peaks in different geographic regions in RS, despite the limited number of investigated samples. Furthermore, the analysis of SARS-CoV-2 diversity from three epidemic peaks provided a profile of circulating lineages in the State and in Brazil and highlighted the massive spread of P.1 and P.1-related lineages in RS. Hence, this study contributes to the genomic surveillance of SARS-CoV-2 and reinforces its importance to guide health authorities in decisions for the control of COVID-19.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.meegid.2021.105144>.

Data deposition

SARS-CoV-2 sequences generated in this study were submitted to GISAID under accession numbers available in Supplementary Data 1.

Funding

This work was supported by the Health Department of Rio Grande do Sul, by the Ministry of Health of Brazil, and by Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS 06/2020 – Grant process 20/2551-0000273-0). ABGV holds a Research Fellowship (PQ2) from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil – Grant process 306369/2019-2), and APMV holds a post-doctoral fellowship from Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA).

Credit Authorship Contribution Statement

Ana Paula Muterle Varela: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft. **Janira Prichula:** Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing – original draft. **Fabiana Quos Mayer:** Conceptualization, Data curation, Investigation, Methodology, Writing – original draft. **Richard Steiner Salvato:** Data curation, Formal analysis, Visualization. **Fernando Hayashi Sant'Anna:** Data curation, Formal analysis. **Tatiana Schäffer Gregianini:** Data curation, Methodology, Resources, Writing – review and editing. **Letícia Garay Martins:** Data acquisition. **Adriana Seixas:** Writing – review and editing. **Ana Beatriz Gorini da Veiga:** Conceptualization, Supervision, Project administration, Resources, Funding acquisition, Writing – review and editing.

Declaration of competing interest

ABGV holds a Research Fellowship (PQ2) from CNPq. Reagents were acquired with funds from grants of Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul. APMV holds a post-doctoral fellowship from Universidade Federal de Ciências da Saúde de Porto Alegre. No other conflicts of interest exist.

Acknowledgments

We thank Dr. Vanessa Mattevi, Dr. Marília Zandoná, and Ludmila Fiorenzano Baethgen for providing technical support. We thank all the authors who have shared genome data on GISAID, and we have included a table (Supplementary Data 1) listing the authors and institutes involved.

References

- Almeida, L.G.P., Lamarca, A.P., Francisco Jr., R.S., Cavalcante, L., Gerber, A.L., Guimarães, A.P.C., Machado, D.T., Alves, C., Mariani, D., Cruz, T.F., Ribeiro, M.S., Carvalho, S., da Silva, F.D., Garcia, M.H.O., Souza, L.M., da Silva, G.D., Ribeiro, C.L.P., Cavalcanti, A.C., Mello, C.M.B., Tanuri, A., Vasconcelos, A.T., 2021. Genomic

- Surveillance of SARS-CoV-2 in the State of Rio de Janeiro, Brazil: A Technical Briefing. <https://virological.org/t/genomic-surveillance-of-sars-cov-2-in-the-state-of-rio-de-janeiro-brazil-technical-briefing/683>.
- Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., Wishart, D.S., 2016. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkw419>.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Genome analysis Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* <https://doi.org/10.1093/bioinformatics/btu170>.
- Candido, D.S., Claro, I.M., de Jesus, J.G., Souza, W.M., Moreira, F.R.R., Dellicour, S., Mellan, T.A., du Plessis, L., Pereira, R.H.M., Sales, F.C.S., Manuli, E.R., Thézé, J., Almeida, L., Menezes, M.T., Voloch, C.M., Fumagalli, M.J., Coletti, T.M., da Silva, C.A.M., Ramundo, M.S., Amorim, M.R., Hoeltgebaum, H.H., Mishra, S., Gill, M.S., Carvalho, L.M., Buss, L.F., Prete, C.A., Ashworth, J., Nakaya, H.I., Peixoto, P.S., Brady, O.J., Nicholls, S.M., Tanuri, A., Rossi, Á.D., Braga, C.K.V., Gerber, A.L., de Guimarães, A.P.C., Gaburo, N., Alencar, C.S., Ferreira, A.C.S., Lima, C.X., Levi, J.E., Granato, C., Ferreira, G.M., Francisco, R.S., Granja, F., Garcia, M.T., Moretti, M.L., Perroud, M.W., Castiñeiras, T.M.P.P., Lazari, C.S., Hill, S.C., de Souza Santos, A.A., Simeoni, C.L., Forato, J., Sposito, A.C., Schreiber, A.Z., Santos, M.N.N., de Sá, C.Z., Souza, R.P., Resende-Moreira, L.C., Teixeira, M.M., Hubner, J., Leme, P.A.F., Moreira, R.G., Nogueira, M.L., Ferguson, N.M., Costa, S.F., Proenca-Modena, J.L., Vasconcelos, A.T.R., Bhatt, S., Lemey, P., Wu, C.H., Rambaut, A., Loman, N.J., Aguiar, R.S., Pybus, O.G., Sabino, E.C., Faria, N.R., 2020. Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science.* <https://doi.org/10.1126/SCIENCE.ABD2161>.
- Coutinho, R.M., Marquitti, F.M.D., Ferreira, L.S., Borges, M.E., Paixão da Silva, R.L., Canton, O., Portella, T.P., Poloni, S., Franco, C., Plucinski, M.M., Lessa, F.C., Moura da Silva, A.A., Kraenkel, R.A., de Sousa Mascena Veras, M.A., Inácio Prado, P., 2021. Model-based estimation of transmissibility and reinfection of SARS-CoV-2 P.1 variant. *medRxiv.* <https://doi.org/10.1101/2021.03.03.21252706>.
- Dejinrattisai, W., Zhou, D., Supasa, P., Liu, C., Mentzer, A.J., Ginn, H.M., Zhao, Y., Duyvesteyn, H.M.E., Tuekprakhon, A., Nutalai, R., Wang, B., López-Camacho, C., Slon-Campos, J., Walter, T.S., Skelly, D., Costa Clemens, S.A., Naveca, F.G., Nascimento, V., Nascimento, F., Fernandes da Costa, C., Resende, P.C., Pauvolid-Correa, A., Siqueira, M.M., Dold, C., Levin, R., Dong, T., Pollard, A.J., Knight, J.C., Crook, D., Lambe, T., Clutterbuck, E., Bibi, S., Flaxman, A., Bittaye, M., Belij-Rammerstorfer, S., Gilbert, S.C., Carroll, M.W., Klenerman, P., Barnes, E., Dunachie, S.J., Paterson, N.G., Williams, M.A., Hall, D.R., Hulsmit, R.J.G., Bowden, T.A., Fry, E.E., Mongkolsapaya, J., Ren, J., Stuart, D.I., Screaton, G.R., 2021. Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell.* <https://doi.org/10.1016/j.cell.2021.03.055>.
- Faria, N.R., Mellan, T.A., Whittaker, C., Claro, I.M., Candido, D., Mishra, S., Crispim, M., Sales, F., Hawryluk, I., McCrone, J.T., Hulsmit, R., Franco, L., Ramundo, M.S., de Jesus, J.G., Andrade, P.S., Coletti, T.M., Ferreira, G.M.S., 2021. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science.* <https://doi.org/10.1126/science.abb2644>.
- Farkas, C., Mella, A., Haigh, J.J., 2020. Large-scale population analysis of SARS-CoV-2 whole genome 1 sequences reveals host-mediated viral evolution with emergence of mutations 2 in the viral Spike protein associated with elevated mortality rates. *medRxiv.* <https://doi.org/10.1101/2020.10.23.20218511>.
- Franceschi, V.B., Caldana, G.D., de Menezes Mayer, A., Cybis, G.B., Neves, C.A.M., Ferrareze, P.A.G., Demoliner, M., de Almeida, P.R., Gularte, J.S., Hansen, A.W., Weber, M.N., Fleck, J.D., Zimmerman, R.A., Kmetzsch, L., Spilki, F.R., Thompson, C.E., 2021. Genomic epidemiology of SARS-CoV-2 in Esteio, Rio Grande do Sul, Brazil. *BMC Genomics.* <https://doi.org/10.1186/s12864-021-07708-w>.
- Francisco, R. da S., Benites, L.F., Lamarca, A.P., de Almeida, L.G.P., Hansen, A.W., Gularte, J.S., Demoliner, M., Gerber, A.L., de Guimarães, A.P., Antunes, A.K.E., Heldt, F.H., Mallmann, L., Hermann, B., Ziulkoski, A.L., Goes, V., Schallenberg, K., Filippi, M., Pereira, F., Weber, M.N., de Almeida, P.R., Fleck, J.D., Vasconcelos, A.T.R., Spilki, F.R., 2021. Pervasive transmission of E484K and emergence of VUI-NP13L with evidence of SARS-CoV-2 co-infection events by two different lineages in Rio Grande do Sul. *Virus Res.* <https://doi.org/10.1016/j.virusres.2021.198345>.
- Fujino, T., Nomoto, H., Kutsuna, S., Ujiie, M., Suzuki, T., Sato, R., Fujimoto, T., Kuroda, M., Wakita, T., Ohmagari, N., 2021. Novel SARS-CoV-2 variant in travelers from Brazil to Japan. *Emerg. Infect. Dis.* <https://doi.org/10.3201/eid2704.210138>.
- García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Götz, S., Tarazona, S., Dopazo, J., Meyer, T.F., Conesa, A., 2012. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics.* <https://doi.org/10.1093/bioinformatics/bts503>.
- Gorini da Veiga, A.B., Kretzmann, N.A., Corrêa, L.T., Goshiyama, A.M., Baccin, T., Ardenghi, P., Matias, F., Gregianini, T.S., Alves d'Azevedo, P., 2012. Viral load and epidemiological profile of patients infected by pandemic influenza A (H1N1) 2009 and seasonal influenza A virus in Southern Brazil. *J. Med. Virol.* <https://doi.org/10.1002/jmv.23198>.
- Gräf, T., Bello, G., Venas, T.M.M., Pereira, E.C., Paixão, A.C.D., Appolinario, L.R., Lopes, R.S., Mendonça, A.C.F., Rocha, A.S.B., Motta, F.C., Gregianini, T.S., Salvato, R.S., Fernandes, S.B., Rovaris, D.B., Cavalcanti, A.C., Leite, A.B., Riediger, I., Debur, M.C., Bernardes, A.F.L., Ribeiro-Rodrigues, R., Grinsztajn, B., Nascimento, V. A., Souza, V.C., Gonçalves, L., Costa, C.F., Mattos, T., Dezordi, F.Z., Wallau, G.L., Naveca, F.G., Delatorre, E., Siqueira, M.M., Resende, P.C., 2021. Identification of SARS-CoV-2 P.1-related lineages in Brazil provides new insights about the mechanisms of emergence of Variants of Concern. *Res. Sq.* <https://doi.org/10.12103/rs.3.rs-580195/v1>.
- Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M., Ludden, C., Reeve, R., Rambaut, A., Peacock, S.J., Robertson, D.L., 2021. SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* <https://doi.org/10.1038/s41579-021-00573-0>.
- Hensley, M.K., Bain, W.G., Jacobs, J., Nambulli, S., Parikh, U., Cillo, A., Staines, B., Heaps, A., Sobolewski, M.D., Rennick, L.J., Macatangay, B.J.C., Klammer-Blain, C., Kitsios, G.D., Methé, B., Somasundaram, A., Bruno, T.C., Cardello, C., Shan, F., Workman, C., Ray, P., Ray, A., Lee, J., Sethi, R., Schwarzmann, W.E., Ladinsky, M.S., Bjorkman, P.J., Vignali, D.A., Duprex, W.P., Agha, M.E., Mellors, J.W., McCormick, K.D., Morris, A., Haidar, G., 2021. Intractable coronavirus disease 2019 (COVID-19) and prolonged severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) replication in a chimeric antigen receptor-modified T-cell therapy recipient: a case study. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciab072>.
- Hoffmann, M., Arora, P., Groß, R., Seidel, A., Hörnich, B.F., Hahn, A.S., Krüger, N., Graichen, L., Hofmann-Winkler, H., Kempf, A., Winkler, M.S., Schulz, S., Jäck, H.M., Jahrsdörfer, B., Schrezenmeier, H., Müller, M., Kleger, A., Münch, J., Pöhlmann, S., 2021. SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies. *Cell.* <https://doi.org/10.1016/j.cell.2021.03.036>.
- Instituto Brasileiro de Geografia e Estatística, IBGE, 2021. <https://www.ibge.gov.br/cidades-e-estados/rs.html>. Accessed date: September 2021.
- Johns Hopkins Coronavirus Resource Center, 2021. <https://coronavirus.jhu.edu/map.html>. Accessed date: September 2021.
- Lai, A., Bergna, A., Caucci, S., Clementi, N., Vicenti, I., Dragoni, F., Cattelan, A.M., Menzo, S., Pan, A., Callegaro, A., Tagliabracchi, A., Caruso, A., Caccuri, F., Ronchiadini, S., Balotta, C., Zazzi, M., Vaccher, E., Clementi, M., Galli, M., Zehender, G., 2020. Molecular tracing of SARS-CoV-2 in Italy in the first three months of the epidemic. *Viruses.* <https://doi.org/10.3390/v12080798>.
- Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* <https://doi.org/10.1093/bioinformatics/btp324>.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. *Bioinformatics.* <https://doi.org/10.1093/bioinformatics/btp352>.
- McCarthy, K.R., Rennick, L.J., Nambulli, S., Robinson-McCarthy, L.R., Bain, W.G., Haidar, G., Paul Duprex, W., 2021. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science.* <https://doi.org/10.1126/science.abb6950>.
- Ministério da Saúde, 2021. <https://covid.saude.gov.br>. Accessed date: September 2021.
- Mir, D., Rego, N., Resende, P.C., Tort, F., Fernández-Calero, T., Noya, V., Brandes, M., Possi, T., Arleo, M., Reyes, N., Victoria, M., Lizasoain, A., Castells, M., Maya, L., Salvo, M., Schäffer Gregianini, T., Mar da Rosa, M.T., Garay Martins, L., Alonso, C., Vega, Y., Salazar, C., Ferrés, I., Smircich, P., Sotelto Silveira, J., Fort, R.S., Mathó, C., Arantes, I., Appolinario, L., Mendonça, A.C., Benítez-Galeano, M.J., Simeos, C., Graña, M., Motta, F., Siqueira, M.M., Bello, G., Colina, R., Spangenberg, L., 2021. Recurrent dissemination of SARS-CoV-2 through the Uruguayan-Brazilian border. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2021.653986>.
- Nagy, A., Pongor, S., Györfy, B., 2021. Different mutations in SARS-CoV-2 associate with severe and mild outcome. *Int. J. Antimicrob. Agents.* <https://doi.org/10.1016/j.ijantimicag.2020.106272>.
- Naveca, F.G., Nascimento, V., de Souza, V.C., Corado, A. de L., Nascimento, F., Silva, G., Costa, Á., Duarte, D., Pessoa, K., Mejía, M., Brandão, M.J., Jesus, M., Gonçalves, L., da Costa, C.F., Sampaio, V., Barros, D., Silva, M., Mattos, T., Pontes, G., Abdalla, L., Santos, J.H., Arantes, I., Dezordi, F.Z., Siqueira, M.M., Wallau, G.L., Resende, P.C., Delatorre, E., Gräf, T., Bello, G., 2021. COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence. *Nat. Med.* <https://doi.org/10.1038/s41591-021-01378-7>.
- Platto, S., Xue, T., Carafoli, E., 2020. COVID19: an announced pandemic. *Cell Death Dis.* <https://doi.org/10.1038/s41419-020-02995-9>.
- Resende, P.C., Delatorre, E., Gräf, T., Mir, D., Motta, F.C., Appolinario, L.R., Paixão, A.C. D. da, Mendonça, A.C. da F., Ogrzewalska, M., Caetano, B., Wallau, G.L., Docena, C., Santos, M.C. dos, de Almeida Ferreira, J., Sousa Junior, E.C., Silva, S.P. da, Fernandes, S.B., Vianna, L.A., Souza, L. da C., Ferro, J.F.G., Nardy, V.B., Santos, C.A., Riediger, I., do Carmo Debur, M., Croda, J., Oliveira, W.K., Abreu, A., Bello, G., Siqueira, M.M., 2021a. Evolutionary dynamics and dissemination pattern of the SARS-CoV-2 lineage B.1.1.33 during the early pandemic phase in Brazil. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2020.615280>.
- Resende, P.C., Naveca, F.G., Lins, R.D., Dezordi, F.Z., Ferraz, M.V.F., Moreira, E.G., Coelho, D.F., Motta, F.C., Paixão, A.C.D., Appolinario, L., Lopes, R.S., da Fonseca Mendonça, A.C., da Rocha, A.S.B., Nascimento, V., Souza, V., Silva, G., Nascimento, F., Gonçalves Lima Neto, L., Riediger, I., Debur, M. do C., Leite, A.B., Mattos, T., da Costa, C.F., Pereira, F.M., Khouri, B., Bernardes, A.F.L., Delatorre, E., Gräf, T., Mendonça Siqueira, M., Bello, G., Wallau, G.L., 2021b. The ongoing evolution of variants of concern and interest of SARS-CoV-2 in Brazil revealed by convergent indels in the amino (N)-terminal domain of the Spike protein. *medRxiv.* <https://doi.org/10.1101/2021.03.19.21253946>.
- Ribes, M., Chaccour, C., Moncunill, G., 2021. Adapt or perish: SARS-CoV-2 antibody escape variants defined by deletions in the Spike N-terminal Domain. *Signal Transduct. Target. Ther.* <https://doi.org/10.1038/s41392-021-00601-8>.
- Salvato, R.S., Gregianini, T.S., Baethgen, L.F., Schiefelbein, S.H., Machado, T.R.M., Becker, I.M., Ramos, R., Piazza, C.F., Nunes, Z.M.A., Campos, A.A.S., Crescente, L.V., Ranieri, T.M.S., Martins, L.G., Silva, E.V. da, Vizeu, S., Vallandro, M.J., Pedroso, E. R., Burille, A., Bastos, C.G.M.B., 2021. Epidemiological investigation reveals local transmission of SARS-CoV-2 lineage P.1 in Southern Brazil. *Rev. Epidemiol. Control. Infecção.* <https://doi.org/10.17058/reci.v11i1.16335>.
- Sant'Anna, F.H., Mutterle Varela, A.P., Prichula, J., Comerlato, J., Comerlato, C.B., Roglio, V.S., Mendes Pereira, G.F., Moreno, F., Seixas, A., Wendland, E.M., 2021. Emergence of the novel SARS-CoV-2 lineage VUI-NP13L and massive spread of P.2 in South Brazil. *Emerg. Microbes Infect.* <https://doi.org/10.1080/22221751.2021.1949948>.

- Secretaria da Saúde do Estado do Rio Grande do Sul, 2021. <https://ti.saude.rs.gov.br/covid19/>. Accessed date: September 2021.
- Voloch, C.M., da Silva Francisco, R., de Almeida, L.G.P., Cardoso, C.C., Brustolini, O.J., Gerber, A.L., Guimarães, A.P. de C., Mariani, D., da Costa, R.M., Ferreira, O.C., Cavalcanti, A.C., Frauches, T.S., de Mello, C.M.B., Leitão, I. de C., Galliez, R.M., Faffe, D.S., Castineiras, T.M.P.P., Tanuri, A., de Vasconcelos, A.T.R., 2021. Genomic characterization of a novel SARS-CoV-2 lineage from Rio de Janeiro, Brazil. *J. Virol.* <https://doi.org/10.1128/jvi.00119-21>.
- Zehender, G., Lai, A., Bergna, A., Meroni, L., Riva, A., Balotta, C., Tarkowski, M., Gabrieli, A., Bernacchia, D., Rusconi, S., Rizzardini, G., Antinori, S., Galli, M., 2020. Genomic characterization and phylogenetic analysis of SARS-COV-2 in Italy. *J. Med. Virol.* <https://doi.org/10.1002/jmv.25794>.
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G.F., Tan, W., 2020. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* <https://doi.org/10.1056/nejmoa2001017>.