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.

Infection, Genetics and Evolution 93 (2021) 104976

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Short communication

Molecular surveillance of the on-going SARS-COV-2 epidemic in Ribeirao Preto City, Brazil

Svetoslav Nanev Slavov [a,1], Marta Giovanetti [b,c,1], Rafael dos Santos Bezerra [a], Vagner Fonseca [c,d], Elaine Vieira Santos [a], Evandra Strazza Rodrigues [a], Talita Adelino [c,e], Joilson Xavier [c], Josiane Serrano Borges [a], Mariane Evaristo [a], Mauricio Teixeira Lima [c], Glauco de Carvalho Pereira [e], Aparecida Yulie Yamamoto [f], Diego Villa Clé [f], Rodrigo Tocantins Calado [a,f], Dimas Tadeu Covas [a,f], Luiz Carlos Junior Alcantara [b,c], Simone Kashima [a,*]

[a] Blood Center of Ribeirão Preto, Faculty of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto, São Paulo, Brazil
[b] Laboratório de Flavivírus, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
[c] Laboratório de Genética Celular e Molecular, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
[d] KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), College of Health Sciences, University of KwaZuluNatal, Durban 4001, South Africa
[e] Laboratório Central de Saúde Pública, Fundação Ezequiel Dias, Belo Horizonte, Brazil
[f] Department of Medical Imaging, Hematology, and Oncology, Ribeirão Preto Medical School, University of São Paulo, Brazil

ARTICLE INFO

Keywords:
SARS-CoV-2
COVID-19
Variants of concern
VOC
P.1
Whole genome
Phylogeny

ABSTRACT

The Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of an unprecedented worldwide pandemic. Brazil demonstrates one of the highest numbers of confirmed SARS-CoV-2 cases, and São Paulo State is the epicenter of the pandemics in the country. Nevertheless, little is known about the SARS-CoV-2 circulation in other cities in the State than São Paulo city. The objective of this study was to analyze phylogenetically SARS-CoV-2 strains circulating in city of Ribeirão Preto at the beginning of the pandemic and during the actual second wave. Twenty-nine nasopharyngeal SARS-CoV-2 RNA positive samples were sequenced by nanopore technology (18 obtained at the initial period of the pandemic and 11 during the second wave) and analyzed them phylogenetically. The performed analysis demonstrated that the majority of the strains obtained in the initial period of the pandemic in Ribeirão Preto belonged mainly to the B1.1.33 lineage (61.1%), but B.1.1 (27.8%) and B.1.1.28 (11.1%) lineages were also identified. In contrast, the second wave strains were composed exclusively by the Brazilian variant of concern (VOC) P.1 (91%) and P.2 (9%) lineages. The obtained phylogenetic results were suggestive of successive SARS-CoV-2 lineage substitution in this Brazilian region by the P.1 VOC. The performed study examines the SARS-CoV-2 genotypes in Ribeirão Preto city via genomic surveillance data. The obtained findings can contribute for continuous long-term genomic surveillance of SARS-CoV-2 due to the accelerated dynamics of viral lineage substitution, predict further waves and examine lineage behavior during SARS-CoV-2 vaccination.

The Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2), causing the coronavirus disease-19 (COVID-19) is a betacoronavirus with a genome of ~30 kb and is currently responsible for one of the largest pandemics registered over the last century since the 1918 flu. The first cases of severe pneumonia caused by SARS-CoV-2 were identified between workers from the Huanan Wholesale market in the city of Wuhan, Hubei Province, China (Andersen et al., 2020). On 30th of January 2020, due to the rapid viral dissemination in almost all countries in the world, the WHO declared public health emergency of international concern and pandemic on 11 March 2020. According to the proposed SARS-CoV-2 lineage nomenclature, two main SARS-CoV-2 lineages (A and B) with multiple sublineages have been already identified (Rambaut et al., 2020).

Currently, Brazil faces one of the highest number of SARS-CoV-2

* Corresponding author at: Laboratory of Molecular Biology, Regional Blood Center of Ribeirão Preto, Avenida Tenente Catão Roxo, 2501, Ribeirão Preto, São Paulo, Brazil.
E-mail address: skashima@hemocentro.fmrp.usp.br (S. Kashima).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.meegid.2021.104976
Received 16 December 2020; Received in revised form 18 June 2021; Accepted 21 June 2021
Available online 24 June 2021

confirmed cases, 17,452,610 (Brazilian Ministry of Health, https://covid.saude.gov.br, June, 15th, 2021) as 20% of them are reported for the São Paulo State which is the epicenter of the disease (3,464,612 confirmed cases by June, 15th, 2021). However, little is still known for the SARS-CoV-2 molecular epidemiology in the inland São Paulo State. Ribeirão Preto city, located in the inland São Paulo State showed one of the highest rates of SARS-CoV-2 confirmed number of cases with an incidence of 6489 confirmed cases per 100,000 inhabitants (https://www.ribeiraopreto.sp.gov.br/portal/pdf/saude17b202106.pdf) and the SARS-CoV-2 molecular epidemiology has not been well studied.

Thus, to better understand the SARS-CoV-2 molecular dynamics in Ribeirão Preto city we sequenced 18 complete SARS-CoV-2 genomes obtained from the initial period of the pandemic (March–April 2020) and 11 one year later during the second ongoing SARS-CoV-2 wave (started by the end of December 2020). The samples were selectively obtained from patients with different clinical outcomes and we performed phylogenetic analysis. The first SARS-CoV-2 autochthonous cases in Ribeirão Preto were registered by our research group in March 2020. The mean age of the patients included in the study was 48.4 years (SD ± 16.5), 16 were male and 13 were female; eight patients eventually succumbed to COVID-19. In the study samples were obtained from patients with different clinical presentations of COVID-19 (mild, medium, severe, and lethal outcome). The study was approved by the local Research Ethics Committee, Ribeirão Preto Medical School, University of São Paulo (Process CAAE number, 38975620.1.1001.5440).

cDNA synthesis reaction was performed on 29 selected (based on cycle threshold values ≤32) samples using SuperScript IV Reverse Transcriptase kit (Invitrogen), following the manufacturer's instructions. Sequencing multiplex PCR was performed following an open access protocol for SARS-COV-2 sequencing (Quick, 2020), using V.1 and V.3 primers pools, designed by ARTIC Network (https://artic.network/ncov-2019). Sequencing libraries were prepared using the Oxford Nanopore Ligation Sequencing Kit (SQK-LSK109) and Native Barcoding Expansion kits (NBD104 and EXP-NBD114) following previously published protocol (Quick et al., 2017). The libraries were loaded on a MinION flow cell (FLO-MIN106) and sequenced within 24 h. Raw files were basecalled using Guppy and barcode demultiplexing was performed using qcat. We used Genome Detective and Coronavirus Typing Tool to obtain consensus sequences by de novo assembling (Cleemput et al., 2020; Vilsker et al., 2019).

Complete SARS-CoV-2 genome sequences were downloaded from GISAID EpiCoV database. Sequences were aligned using MAFFT (FF-NS-2 algorithm) following default parameters (Katoh et al., 2019). The alignment was manually inspected to remove artefacts using Aliview software (Larsson, 2014). A Maximum Likelihood (ML) phylogeny was inferred on a dataset containing the 29 new sequences plus other 3873 reference sequences deposited in GISAID up to 15 April 2021, using IQ-TREE (version 2.0.5) under the GTR + G4 + F model according to Bayesian Information Criterion (BIC) indicated by the Model Finder in IQ-TREE (Nguyen et al., 2015). An ultrafast bootstrap approximation with 1000 replicates was used to assess branch support. The reference SARS-CoV-2 strains composing the phylogenetic tree were obtained from an available dataset from the Nextclade (https://clades.nextstrain.org).

The performed phylogenetic analysis demonstrated that the SARS-CoV-2 genomes obtained from the initial stages of the pandemic (March–April 2020) belonged mainly to the B.1.1.33 lineage ($n = 11/18$, 61.1%) followed by B.1.1 ($n = 5/18$; 27.8%) and B.1.1.28 ($n = 2/18$;

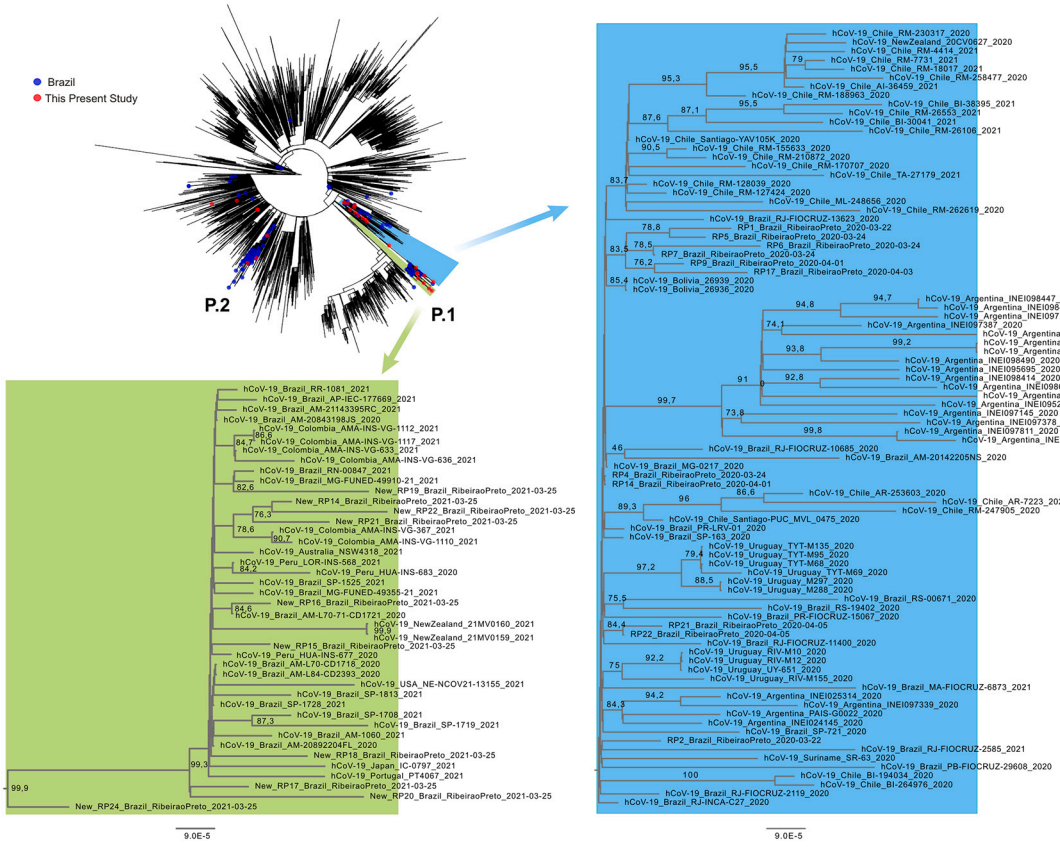

Fig. 1. Phylogenetic analysis of SARS-CoV-2 genomes obtained in Ribeirão Preto city, São Paulo state during the initial period of the pandemic in the region and the second SARS-CoV-2 wave. Only complete SARS-CoV-2 genomes obtained from GISAID by April 15th, 2021 were used for tracing the phylogenetic history. The nucleotide substitution model used was GTR + G4 + F for tree reconstruction, which was chosen by BIC (Bayesian Information Criterion) statistic model, utilizing 1000 ultrafast bootstrap replicates for statistical significance. Only values of above 75% were demonstrated on important tree branches. The phylogenetic tree was constructed using the IQtree software v.16.12, applying the maximum likelihood approach.

11.1%) lineages. On the contrary, all genomes analyzed during the second SARS-CoV-2 wave in Ribeirão Preto (March 2021) belonged to the Brazilian variants: P.1 variant of concern (VOC) ($n = 10/11$, 91%) and P.2 (n = 1/11; 9%) (Fig. 1). In the dendrogram, the samples from this study classified as P.1 were randomly distributed along the cluster with other P.1 strains circulating in Brazil, which is an indication for the wide dispersion of this VOC. Taken together, the obtained data suggest that the SARS-CoV-2 outbreak in Ribeirão Prato is dynamically evolving since the first SARS-CoV-2 introduction in the region, when the initial SARS-CoV-2 lineages corresponding to B.1.1.28 and B.1.1.33 were completely substituted by the P.1 lineage. This corresponds to the actual epidemiological situation in Ribeirão Preto and in Brazil, which is related to rise of the newly confirmed cases and high morbidity and mortality.

We evaluated the molecular evolution of SARS-CoV-2 lineages in the city of Ribeirão Preto from the initial period of the SARS-CoV-2 pandemic and the most recent wave related to high rise in the number of confirmed cases. Our analysis showed that at the initial period of pandemic, SARS-CoV-2 strains were taxonomically classified as B.1.1.28 and B.1.1.33 lineages, which is reported by other studies (Candido et al., 2020). Nevertheless, the majority of the recently analyzed samples were classified as belonging to the Brazilian P.1 VOC which is currently dominating the epidemiological scenario in Brazil. Such a result shows, despite of the small number of analyzed samples, that in our region the P.1 VOC has largely substituted in circulation the initially identified B1.1.33 and B1.1.28 lineages (March–April 2020). A similar molecular epidemiological pattern has also been observed in a study performed in the Brazilian city of Manaus, where the P.1 VOC was initially identified. In this location, the rapid spread and faster molecular evolution of the P.1 VOC compared to the initial strains led to unprecedented rise in the SARS-CoV-2 cases between November–December 2020 (Faria et al., 2021; Naveca et al., 2021) despite of the high SARS-CoV-2 seroprevalence. Similar situation was also observed regarding other VOCs like B.1.1.7 (Volz et al., 2021) and B.1.351 (Tegally et al., 2021) lineages. The continuous SARS-CoV-2 transmission creates favorable conditions for the emergence of viral variants which show rapid displacement over the non-VOC lineages, which is a result of increased transmissibility (Tegally et al., 2021; Volz et al., 2021). An interesting observation in support of this is the monophyletic separation in the performed phylogenetic analysis of the P.1 cluster composed almost exclusively of Brazilian isolates which shows that this lineage has emerged in Brazil and demonstrates sustained transmission which shapes the current Brazilian SARS-CoV-2 scenario.

Nevertheless, our study shows a small part of the overall burden of the P.1 VOC dissemination in this Brazilian region and therefore more studies including analysis of a higher number of SARS-CoV-2 isolates are necessary to more comprehensively understand the evolution and molecular epidemiology of SARS-CoV-2 in this region especially the origin of variants like P.1. The pathogenesis of SARS-CoV-2 severe disease is still unknown. The random distribution of the sequenced isolates throughout the reconstructed phylogenetic tree in our study suggests that host factors rather than viral genetic variations are more relevant to determine disease severity. Previous studies suggest that genetic determinants and predictors of host immunity are related to the susceptibility to infection and the COVID-19 clinical outcome (Ramlall et al., 2020).

In conclusion, this study examines the SARS-CoV-2 molecular evolution in Ribeirão Preto via SARS-CoV-2 genome surveillance data. These findings can contribute for the long-term genomic surveillance of SARS-CoV-2 in the examined region as well as the genomic evaluation of the circulating strains in further outbreaks and vaccine policy applications.

Funding

This work was supported by Centro de Terapia Celular (CTC) - FAPESP (2013/08135-2; 2018/15826-5; 17/26950-6), INCTC, (465539/2014-9), FINEP (FMUSP, N207.234), FUNDHERP. We are grateful for the support provided by the personnel from the Central Public Health Laboratory/Octavio Magalhaes Institute (IOM) of the Ezequiel Dias Foundation (FUNED). This work was supported by the Pan-American Health Organization (IOC-007-FEX-19-2-2-30). MG receives grant from the Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

Data availability statement

SARS-CoV-2 genome sequences generated in this study have been deposited in the GISAID platform (https://www.gisai-d.org) under the following accession numbers: EPI_ISL_613563; EPI_ISL_613564; EPI_ISL_613707, EPI_ISL_613708; EPI_ISL_613709; EPI_ISL_613711; EPI_ISL_613951; EPI_ISL_613952; EPI_ISL_613954; EPI_ISL_613956; EPI_ISL_613961; EPI_ISL_613962; EPI_ISL_613963; EPI_ISL_613964; EPI_ISL_613965; EPI_ISL_614011; EPI_ISL_614155; EPI_ISL_614156; EPI_ISL_1786560 - EPI_ISL_1786570.

Declaration of Competing Interest

No potential conflict of interest was reported by the authors.

References

Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C., Garry, R.F., 2020. The proximal origin of SARS-CoV-2. Nat. Med. 26 (4), 450–452.
Candido, D.S., Claro, I.M., de Jesus, J.G., Souza, W.M., Moreira, F.R.R., Dellicour, S., et al., 2020. Evolution and epidemic spread of SARS-CoV-2 in Brazil. Science 369, 1255–1260.
Cleemput, S., Dumon, W., Fonseca, V., Abdool Karim, W., Giovanetti, M., Alcantara, L.C., Deforche, K., de Oliveira, T., 2020. Genome Detective Coronavirus Typing Tool for rapid identification and characterization of novel coronavirus genomes. Bioinformatics 36 (11), 3552–3555.
Faria, N.R., Mellan, T.A., Whittaker, C., Claro, I.M., Candido, D.D.S., Mishra, S., et al., 2021. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science. https://doi.org/10.1126/science.abh2644.
Katoh, K., Rozewicki, J., Yamada, K.D., 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. 20, 1160–1166.
Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30 (22), 3276–3278.
Naveca, F.G., Nascimento, V., de Souza, V.C., Corado, A.L., Nascimento, F., Silva, G., et al., 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. Epub ahead of print.
Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32 (1), 268–274.
Quick, Josh, 2020. nCoV-2019 Sequencing Protocol, 3. https://doi.org/10.17504/protocols.io.bbmuik6w.
Quick, J., Grubaugh, N.D., Pullan, S.T., Claro, I.M., Smith, A.D., Gangavarapu, K., et al., 2017. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat. Protoc. 12, 1261–1276.
Rambaut, A., Holmes, E.C., O'Toole, Á., Hill, V., McCrone, J.T., Ruis, C., et al., 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat. Microbiol. 5 (11), 1403–1407.
Ramlall, V., Thangaraj, P.M., Meydan, C., Foox, J., Butler, D., Kim, J., et al., 2020. Immune complement and coagulation dysfunction in adverse outcomes of SARS-CoV-2 infection. Nat. Med. 26 (10), 1609–1615.
Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., et al., 2021. Detection of a SARS-CoV-2 variant of concern in South Africa. Nature 592 (7854), 438–443.
Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, K., et al., 2019. Genome detective: an automated system for virus identification from high-throughput sequencing data. Bioinformatics 35 (5), 871–873.
Volz, E., Mishra, S., Chand, M., Barrett, J.C., Johnson, R., Geidelberg, L., et al., 2021. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. Nature. https://doi.org/10.1038/s41586-021-03470-x.