

# Genomic and phylogenetic characterisation of an imported case of SARS-CoV-2 in Amazonas State, Brazil

Valdinete Alves do Nascimento<sup>1,2</sup>, André de Lima Guerra Corado<sup>1,2</sup>,  
Fernanda Oliveira do Nascimento<sup>1,3</sup>, Ágatha Kelly Araújo da Costa<sup>1,3</sup>,  
Debora Camila Gomes Duarte<sup>1</sup>, Sérgio Luiz Bessa Luz<sup>1,3</sup>, Luciana Mara Fé Gonçalves<sup>4</sup>,  
Michele Silva de Jesus<sup>1</sup>, Cristiano Fernandes da Costa<sup>4</sup>,  
Edson Delatorre<sup>5</sup>, Felipe Gomes Naveca<sup>1,2,3,6/+</sup>

<sup>1</sup>Fundação Oswaldo Cruz-Fiocruz, Instituto Leônidas e Maria Deane, Manaus, AM, Brasil

<sup>2</sup>Fundação Oswaldo Cruz-Fiocruz, Instituto Oswaldo Cruz, Programa de Pós-Graduação em Biologia Celular e Molecular, Rio de Janeiro, RJ, Brasil

<sup>3</sup>Fundação Oswaldo Cruz-Fiocruz, Instituto Leônidas e Maria Deane, Programa de Pós-Graduação em Biologia da Interação Patógeno-Hospedeiro, Manaus, AM, Brasil

<sup>4</sup>Fundação de Vigilância em Saúde do Amazonas, Manaus, AM, Brasil

<sup>5</sup>Universidade Federal do Espírito Santo, Centro de Ciências Exatas, Naturais e da Saúde, Departamento de Biologia, Vitória, ES, Brasil

<sup>6</sup>Rede Genômica de Vigilância em Saúde do Estado do Amazonas, Manaus, AM, Brasil

A new coronavirus [severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)] is currently causing a life-threatening pandemic. In this study, we report the complete genome sequencing and genetic characterisation of a SARS-CoV-2 detected in Manaus, Amazonas, Brazil, and the protocol we designed to generate high-quality SARS-CoV-2 full genome data. The isolate was obtained from an asymptomatic carrier returning from Madrid, Spain. Nucleotide sequence analysis showed a total of nine mutations in comparison with the original human case in Wuhan, China, and support this case as belonging to the recently proposed lineage A.2. Phylogeographic analysis further confirmed the likely European origin of this case. To our knowledge, this is the first SARS-CoV-2 genome obtained from the North Brazilian Region. We believe that the information generated in this study may contribute to the ongoing efforts toward the SARS-CoV-2 emergence.

Key words: coronavirus - SARS-CoV-2 - COVID-19 - Brazil - Amazon Region - genome

The coronavirus disease 2019 (COVID-19) is caused by infection with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that was identified for the first time in patients with pneumonia in Wuhan, China.<sup>(1,2)</sup> Since its discovery at the end of 2019, SARS-CoV-2 transmission has been documented as being person-to-person, causing from an asymptomatic status, that may also generate transmission clusters, to a symptomatic disease with the typical symptoms manifested as fever, dry cough, myalgia, fatigue, dyspnea, diarrhea, and nausea.<sup>(3,4)</sup> In the majority of the patients, the clinical outcome is a mild disease, although Wu and McGoogan described Chinese patients that developed a severe outcome (16%) or a critical condition (4%). Severe or criti-

cal outcomes usually occur in patients with comorbidities, and the disease can progress, presenting arrhythmia and shock, that could evolve to death.<sup>(5,6)</sup>

The SARS-CoV-2 belongs to family *Coronaviridae*, genus *Betacoronavirus*. In 2002 and 2012, two outbreaks occurred caused by new coronaviruses, SARS-CoV and MERS-CoV, with lethality ranging from nine to 33%, respectively.<sup>(7,8,9)</sup> The coronaviruses are enveloped viruses, with 80 to 120 nm in diameter. Three proteins compose the virion surface: spike (S), membrane (M), and small membrane protein (E), giving the virus a crown-like visual on electron micrograph.<sup>(10,11)</sup> This viral family has one of the largest RNA genome of all other RNA viruses, with a single non-segmented positive-stranded RNA of approximately 30 kb.<sup>(12)</sup>

In Brazil, the first recorded case of SARS-CoV-2 infection occurred in the end of February, in the city of São Paulo, followed by cases in the Northeast (Bahia), Central-West (Brasília) and South (Rio Grande do Sul) regions.<sup>(13)</sup> The northern of Brazil was the last region of the country to detect SARS-CoV-2 in its population. In March 13th, the first case in the State of Amazonas was detected in a woman that traveled to England and return to the Amazonas capital, Manaus.<sup>(14)</sup> After three days, the second case of SARS-CoV-2 was detected in Manaus and characterised in the present study.

One 56-years-old man returning from Madrid, Spain, arrived asymptomatic in Manaus, Amazonas State, Brazil on 15-Mar-2020. At that time, a massive outbreak of COVID-19 was already established in several European countries, suggesting that travelers should be kept in

doi: 10.1590/0074-0276200310

Financial support: CNPq, CAPES, MS-DECIT, FIOCRUZ, FAPEAM - REGESAM.

FGN is funded by FAPEAM (PCTI-EmergeSaúde/AM call N°005/2020 and Rede Genômica de Vigilância em Saúde - REGESAM), CNPq (Grants 440856/2016-7 and 403276/2020-9), CAPES (Grants 88881.130825/2016-01 and 88887.130823/2016-00), Inova Fiocruz/FIOCRUZ (Grant VPPCB-007-FIO-18-2-30 - Geração de conhecimento).

+ Corresponding author: felipe.naveca@fiocruz.br

https://orcid.org/0000-0002-2888-1060

Received 14 June 2020

Accepted 25 August 2020



quarantine at arrival. Even without the symptoms of a respiratory infection, nasal and oropharyngeal swabs were collected for SARS-CoV-2 testing as a routine established for respiratory virus surveillance at Instituto Leônidas e Maria Deane (ILMD) - Fiocruz, Amazonas State, Brazil, since 2019. The swabs were combined and submitted to total nucleic acid extraction with a commercial kit (Biogene, Recife, PE, Brazil) and immediately evaluated using the reverse transcription real-time polymerase chain reaction (RT-qPCR) protocol developed by the US Centers for Disease Control and Prevention (CDC/USA) (On 15-Mar-2020, CDC updated the RT-qPCR protocol removing the N3 target). This RT-qPCR assay employs different primers and probes sets aiming three regions of the SARS-CoV-2 nucleocapsid (N) gene (<https://www.fda.gov/media/134922/download>), and the human RNase P as an internal control.

The analysed sample tested positive for SARS-CoV-2 with Cts values of 14.43 (N1), 15.39 (N2) and 15.33 (N3). Thus, we immediately generated cDNA with random primers and Superscript IV reverse transcriptase (ThermoFisher Scientific, Waltham, MA, United States). We had previously designed a PCR scheme to amplify the entire genome of the SARS-CoV-2 based on an alignment of all complete genome sequences available on GenBank at 03/03/2020. Conserved regions were chosen for primer design with Primer3 v2.3.7. embedded in Geneious Software 10.2.6,<sup>(15)</sup> spanning around 2.2Kb and with an overlap region between 131 and 225 bp. This PCR scheme resulted in 15 amplicons with further details presented in the **Supplementary data (Table I)**.

The SARS-CoV-2 whole genome was amplified with Platinum SuperFi II Green PCR master mix (ThermoFisher Scientific) using the 15 primers sets in individual reactions. Each amplicon was then visualised as a unique and intense DNA fragment on agarose gel electrophoresis stained with GelRed (Biotium, Hayward, CA, United States). The PCR amplicons were precipitated with molecular biology grade Polietilenoglicol (PEG) 8,000 (Promega, Madison, WI, United States) and then

resuspended in nuclease-free water. After one-hour incubation at 37°C, all amplicons were quantified in ng/μL using Qubit 2.0 and the dsDNA HS assay kit (ThermoFisher Scientific). Finally, the number of DNA copies in each purified amplicon was estimated with ENDMEMO (<http://www.endmemo.com/bio/dnacopynum.php>), normalised, and pooled. A single library was constructed using the Nextera DNA Flex Library Prep and clustered with MiSeq Reagent Micro Kit v2 (300-cycles), following the manufacturer's protocols. Nucleotide sequencing was performed in the MiSeq platform (Illumina, San Diego, CA, United States), installed at ILMD, in a paired-end run (2x150 cycles).

A total of 10,946,898 reads were trimmed for quality and adapters using BBDOUK v37.25, embedded in Geneious software. Thus, 8,362,418 reads were mapped to the SARS-CoV-2 NCBI Reference Sequence NC\_045512.2 using Geneious map-to-reference tool. The BR\_AM\_ILMD\_20140001 final consensus genome sequence contains 29,789 nucleotides, with no gaps, a Q40 score of 100%, with no undetermined "N" bases and high average coverage (> 34,000X). To avoid any primers bias, we removed both primer binding sites at the 5' and 3' ends. Thus, our final sequence represents the positions between nucleotides 47 and 29,835 or 99.6% of the NCBI RefSeq previously mentioned.

We aligned the BR\_AM\_ILMD\_20140001 genomic sequence with the SARS-CoV-2 NCBI Reference Sequence NC\_045512 using MAFFT v7.388<sup>(16)</sup> to investigate any mutations throughout genome. A total of nine mutations were observed at nucleotide positions 8,782 (C to T); 9,477 (T to A); 12,781 (C to T); 14,805 (C to T); 25,979 (G to T); 26,642 (C to T); 28,144 (T to C); 28,657 (C to T) and 28,863 (C to T), with four of these leading to residues substitution in the deduced protein sequences (Table).

To put the BR\_AM\_ILMD\_20140001 genome in a global context, we aligned the new genome to a pool of all SARS-CoV-2 genomes with at least 25,000 nucleotides available at GISAID database<sup>(17)</sup> on March 31st, 2020 using MAFFT. We adopted a subsampling strategy to reduce

TABLE  
Differences observed between sample BR\_AM\_ILMD\_20140001 and the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) prototype

Genome position	8782	9477	12781	14805	25979	26642	28144	28657	28863
NC_045512	C	T	C	C	G	C	T	C	C
BR_AM_ILMD	T	A	T	T	T	T	C	T	T
Codon position	AGC	TTT	TAC	TAC	GGA	GCC	TTA	GAC	TCA
	AGT	TAT	TAT	TAT	GTA	GCT	TCA	GAT	TTA
Mutation type	s	ns	s	s	ns	s	ns	s	ns
Protein	nsp4	nsp4	nsp9	RdRp	ORF3a ptn	M glycoptn	ORF8 ptn	npptn	npptn
Residue	Ser2839	Phe3071Tyr	Tyr4172	Tyr4847	Gly196Val	Ala40	Leu84Ser	Asp128	Ser197Leu

Nucleotides substituted in each codon are represented in bold. s: silent mutation; ns: non-silent mutation; nsp: non-structural protein; RdRp: RNA-dependent RNA-polymerase; ORF3a ptn: ORF3a protein; M glycoptn: M glycoprotein; ORF8 ptn: ORF8 protein; npptn: nucleocapsid phosphoprotein.

computation time, selecting subsets of 15-20 sequences retaining the most viral diversity from each country using the CD-HIT program.<sup>(18)</sup> This subsampling approach resulted in a final sequence dataset with 490 sequences from 53 countries [Supplementary data (Table II)].

Complete coding sequences (CDS) were subjected to maximum likelihood (ML) phylogenetic reconstruction with PhyML v3.0,<sup>(19)</sup> under the HKY+Γ4 nucleotide substitution model. According to the ML phylogenetic tree, the BR\_AM\_ILMD\_20140001 sequence belonged to the lineage A<sup>(20)</sup> (Fig. 1A) clustering within a monophyletic cluster (aLRT = 1.00) comprising sequences from Spain, Chile, France, Greece, Georgia, Netherlands, Senegal (Fig. 1B). The pangolin web application (pangolin.cog-uk.io) further assigned the sequences from this cluster to the A.2 lineage (Fig. 1B).

After temporal validation with Tempest<sup>(21)</sup> (Fig. 2A), we conducted a Bayesian discrete spatiotemporal analysis with all lineage A sequences (Fig. 2) using the BEAST v1.10 package,<sup>(22)</sup> applying the strict molecular clock and the parametric exponential coalescent models. The strict molecular clock model was selected over the relaxed uncorrelated one,<sup>(23)</sup> after marginal likelihood estimation using path sampling and stepping-stone sampling methods<sup>(24)</sup> [Supplementary data (Table III)]. The analysis under the uncorrelated relaxed molecular clock

model, however, resulted in estimates similar to those of the strict clock model (data not shown). We estimated that the most recent common ancestor (MRCA) of lineage A.2 originated in Spain (posterior state probability, PSP = 0.99) in the beginning of February 2020 (95% highest posterior density, HPD = 2nd - 22nd Feb 2020). The phylogeographic analysis also pointed out that the BR\_AM\_ILMD\_20140001 was most probably introduced from Spain (PSP = 0.43) (Fig. 2B).

Like other viral infections, the infection by SARS-CoV-2 can be asymptomatic and this fact has been reported in different studies.<sup>(25,26,27)</sup> It is noteworthy that according to the definition of suspected case adopted in Brazil at the time that we investigated this asymptomatic carrier, he would not be included in SARS-CoV-2 testing, despite returning from an area with active transmission like Madrid, Spain.<sup>(28)</sup> Until 24-Mar-2020, when we deposited the BR\_AM\_ILMD\_20140001 genome information at <https://www.gisaid.org>, there were only 17 Brazilian SARS-CoV-2 complete genome sequences, and the sequence reported in the present work is the first complete genome from the northern Brazilian region.

Several laboratories over the world are now sequencing thousands of SARS-CoV-2 genomes, which is undoubtedly the most notable effort of viral sequencing in human history. Like any other virus, the new coronavi-

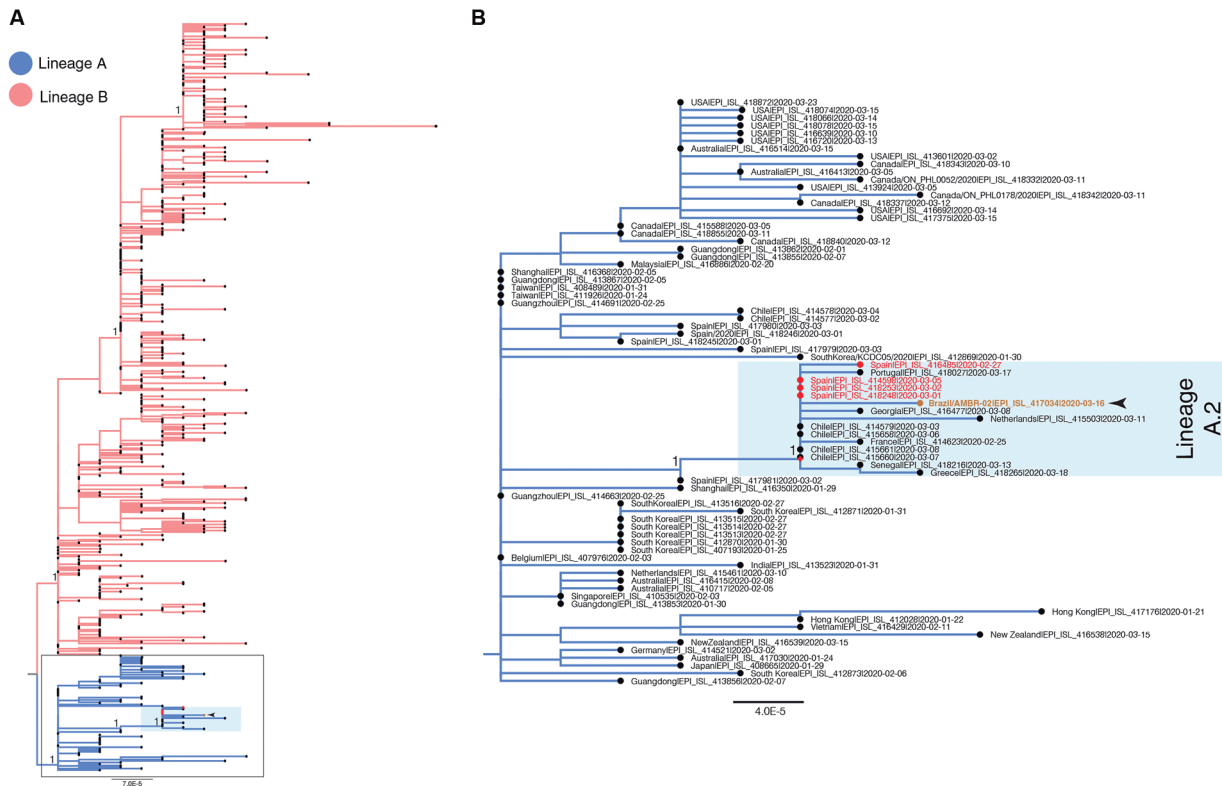


Fig. 1: maximum-likelihood (ML) phylogeny of subsampled severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genomes. (A) ML tree rooted on the branch separating lineages A and B sequences colored following the legend. (B) A close view of the lineage A highlighting the A.2 lineage (light blue box) that comprises the BR\_AM\_ILMD\_20140001 strain (indicated with an arrow). The nodes representing the most recent common ancestor (MRCA) of the lineage A.2 is indicated with a red diamond. In both trees, tips representing the Spanish strains inside lineage A.2 are coloured red and the scale bar represents nucleotide substitutions per site.

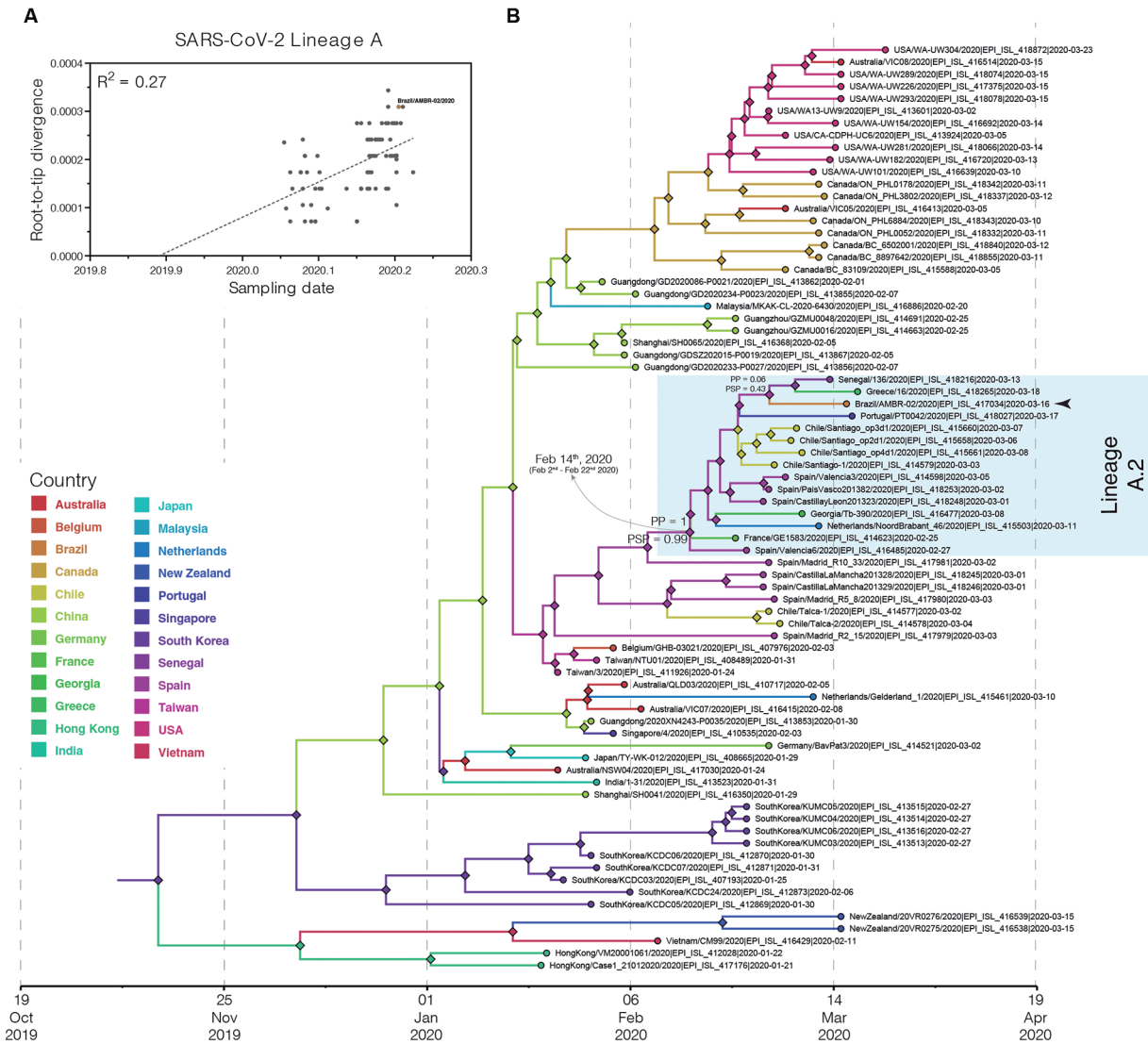


Fig. 2: phylogeography of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) lineage A. (A) Temporal signal analysis correlating the sampling date of each sequence and its genetic distance from the root of a maximum likelihood phylogeny. (B) Time-scaled Bayesian phylogeographic maximum clade credibility (MCC) tree of the SARS-CoV-2 complete coding sequences (CDS) classified as lineage A. The branch's colors represent the most probable location of their descendent nodes (diamonds) as indicated at the legend. Branch support are indicated only at key nodes [posterior (PP) and posterior state probability (PSP)]. The lineage A.2 is highlighted with a light blue box. All horizontal branch lengths are drawn to a scale of years.

rus is continuously evolving as more hosts, humans or animals, are getting infected. Thus, it is of paramount importance to generate and share high-quality full viral genomes from different regions over the world to better understand the SARS-CoV-2 evolution. The information related to the viral evolution is not only necessary for molecular epidemiology studies, but also to monitor if the newly identified mutations are linked to different clinical presentations or may drive into false-negative results when performing nucleic acid amplification assays, like real-time PCR.

Therefore, in this work, we aimed to describe and characterise the complete genome of the SARS-CoV-2 obtained from an asymptomatic carrier returning from Madrid, Spain. To achieve this goal, we decided to use

a nucleotide sequencing strategy where firstly all the 15 amplicons, encompassing the entire SARS-CoV-2 genome, were confirmed by agarose gel electrophoresis. Subsequently, each amplicon was quantified and normalised in order to prevent that one region could be overrepresented during sequencing. In order to make our approach more straightforward, we decided to evaluate if longer amplicons could be generated in a very similar way. We were successful in generating amplicons around 6 Kb, with a minimum overlap of 131 bp, reducing the number of PCR reactions to 5 instead of 15 (data not showed). The conditions to amplify the 6 Kb amplicons followed the manufacturer's recommendations, with the primers pairs details presented in the **Supplementary data (Table I)** of this manuscript. Of

note, we observed that only samples with high viral load (e.g. Ct lower than 21) were fully amplified using the 6Kb protocol. We believe that this approach may be exciting not only to reduce the current protocol costs, but also for those interested in using long reads sequencing technologies like PacBio SMS and nanopore.

Recently, Rambaut and colleagues proposed a rational and dynamic virus classification for SARS-CoV-2 genomes based on a phylogenetic framework.<sup>(20)</sup> Using this approach authors identified at the root of the phylogeny of SARS-CoV-2 two lineages that were simply denoted as lineages A and B. In our analysis, while all Brazilian SARS-CoV-2 sequences belonged to the lineage B, mostly from the B.1 lineage, the BR\_AM\_ILMD\_20140001 genome clustered within the lineage A.2, indicating that BR\_AM\_ILMD\_20140001 strain belongs to a distinct transmission cluster than the other full Brazilian genomes reported until March 31st, 2020. Originally, the lineage B.1 predominated in Europe and North America,<sup>(17,20)</sup> and currently, constitutes the most prevalent SARS-CoV-2 variant circulating in Brazil.<sup>(29)</sup> The lineage A.2 constitutes a predominantly Spanish lineage, found in at least 25 countries<sup>(20)</sup> and our phylogeographic analysis corroborates its origins and reinforces the importation scenario from Spain to Brazil.

In this study, we report and characterise the first SARS-CoV-2 genome obtained from an infected subject in the Brazilian North Region. Since this case was an asymptomatic carrier, it is not easy to suggest when infection has occurred. However, our phylogeographic analysis strongly indicates this individual was infected in Spain. Finally, we would like to emphasise that more fully genomes studies of the SARS-CoV-2 are necessary to better understand the evolution of this emerging life-threatening virus and the information of the nucleotide sequence described here may contribute to future molecular epidemiological studies in Brazil. In this sense, the protocol that we described in the present study may be useful to aid other researchers to generate other high-quality SARS-CoV-2 genomes.

*Nucleotide sequence accession number* - The complete genome sequence of the BR\_AM\_ILMD\_20140001 isolate is available in GISAID since March 24, 2020, under the ID number EPI\_ISL\_417034.

#### ACKNOWLEDGEMENTS

To the following authors from the originating laboratories responsible for obtaining the specimens and the submitting laboratories where genetic sequence data were generated and shared via the GISAID Initiative, on which this research is based. All GISAID accession numbers and data contributors are described in the Supplementary data (Table IV). The authors thank the Program for Technological Development in Tools for Health - FIOCRUZ - for use of nucleotide sequencing facilities at ILMD - Fiocruz Amazônia.

#### AUTHORS' CONTRIBUTION

VAN, SLBL, CFC and FGN conceived the study; FGN designed the study protocol; VAN, ALGC, FON, AKAC, DCGD, LMFG, MSJ and FGN performed the molecular tests; VAN, ED and FGN performed the analysis and interpretation

of these data; FON and MSJ collected biological sample; VAN, ALGC, AKAC, FON, DCGD, LMFG, ED and FGN wrote the manuscript; SLBL, CFC, ED and FGN critically revised the manuscript for intellectual content; FGN financed the study. All authors read and approved the final manuscript. . The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### REFERENCES

- Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579(7798): 270-3.
- Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020; 579(7798): 265-9.
- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*. 2020; 382(13): 1199-207.
- Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*. 2020; 368(6490): 489-93.
- Wu Z, McGoogan JM. Characteristics of and important lessons from the Coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA*. 2020; 323(13): 1239-42.
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020; 395(10223): 497-506.
- Peiris JSM, Lai ST, Poon LLM, Guan Y, Yam LYC, Lim W, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet*. 2003; 61(9366): 1319-25.
- Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med*. 2012; 367(19): 1814-20.
- Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol*. 2016; 24(6): 490-502.
- Bond CW, Leibowitz JL, Robb JA. Pathogenic murine coronaviruses. II. Characterization of virus-specific proteins of murine coronaviruses JHMV and A59V. *Virology*. 1979; 94(2): 371-84.
- Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol Mol Biol Rev*. 2005; 69(4): 635-64.
- Lai MMC, Cavanaght D. The molecular biology of coronaviruses. *Adv Viral Res*. 1997; 48: 1-100.
- MS/SVE - Ministério da Saúde/Secretaria de Vigilância Epidemiológica. Boletim epidemiológico, boletim especial: doença pelo coronavírus Semana epidemiológica 15 (05-10/04), 6 de Abril de 2020. Brasília: Ministério da Saúde; 2020. Available from: <https://www.saude.gov.br/images/pdf/2020/Abril/06/2020-04-06-BE7-Boletim-Especial-do-COE-Atualizacao-da-Avaliacao-de-Risco.pdf>.
- FVS/AM - Fundação de Vigilância em Saúde do Amazonas. Situação epidemiológica de COVID-19 e da síndrome respiratória aguda grave no Estado do Amazonas, Ano 01, Número 1, 27 de Março de 2020. Manaus: Amazonas; 2020. Available from: [http://www.fvs.am.gov.br/media/publicacao/Boletim\\_Situacao\\_Epidemiologica\\_de\\_COVID-19\\_e\\_da\\_S%C3%ADndrome\\_Respiratoria\\_Aguda\\_g9E6Skz.pdf](http://www.fvs.am.gov.br/media/publicacao/Boletim_Situacao_Epidemiologica_de_COVID-19_e_da_S%C3%ADndrome_Respiratoria_Aguda_g9E6Skz.pdf).
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012; 28(12): 1647-9.

16. Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 2002; 30(14): 3059-66.
17. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* 2017; 22(13): 957.
18. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012; 28(23): 3150-2.
19. Guindon S, Lethiec F, Duroux P, Gascuel O. PHYML Online - A web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* 2005; 33(Suppl. 2): 557-9.
20. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol.* 2020; doi: 10.1038/s41564-020-0770-5. Epub ahead of print.
21. Rambaut A, Lam TT, Carvalho LM, Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* 2016; 2(1): 1-7.
22. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 2018; 4(1): 15.
23. Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 2006; 4(5): e88. doi: 10.1371/journal.pbio.0040088.
24. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol Biol Evol.* 2012; 29(9): 2157-67.
25. Pan X, Chen D, Xia Y, Wu X, Li T, Ou X, et al. Asymptomatic cases in a family cluster with SARS-CoV-2 infection. *Lancet Infect Dis.* 2020; 20(4): 410-1.
26. Hoehl S, Rabenau H, Berger A, Kortenbusch M, Cinatl J, Bojkova D, et al. Evidence of SARS-CoV-2 infection in returning travelers from Wuhan, China. *N Engl J Med.* 2020; 382(13): 1278-80.
27. Bai Y, Yao L, Wei T, Tian F, Jin D-Y, Chen L, et al. Presumed asymptomatic carrier transmission of COVID-19. *JAMA.* 2020; 323(14): 1406-7.
28. MS/SVE - Ministério da Saúde/Secretaria de Vigilância Epidemiológica. Boletim epidemiológico, doença pelo coronavírus 2019, 14 de Março de 2020. Brasília: Ministério da Saúde; 2020. Available from: <https://www.saude.gov.br/images/pdf/2020/marco/24/03--ERRATA---Boletim-Epidemiologico-05.pdf>.
29. Resende PC, Delatorre E, Gräf T, Mir D, Motta FC, Appolinario LR, et al. Genomic surveillance of SARS-CoV-2 reveals community transmission of a major lineage during the early pandemic phase in Brazil. Preprint. Available from: <https://www.biorxiv.org/content/10.1101/2020.06.17.158006v1.full.pdf>.