

1 **Article Summary Line:** Active genomic surveillance of SARS-CoV-2 suspected cases from
2 recent travelers reveals the circulation of the P1 variant of concern in Bahia state, Northeast
3 Brazil.

4
5 **Running Title:** Genomic surveillance of P.1 variant in Bahia state

6
7 **Keywords:** SARS-CoV-2, P.1 variant, Genomic surveillance, Brazil

8
9 **Title:** Early genomic detection of SARS-CoV-2 P.1 variant in Northeast Brazil

10
11 **Authors:** Stephane Tosta*, Marta Giovanetti*, Vanessa Brandão Nardy, Luciana Reboredo de
12 Oliveira da Silva, Marcela Kelly Astete Gómez, Jaqueline Gomes Lima, Cristiane Wanderley
13 Cardoso, Tarcisio Oliveira Silva, Marcia São Pedro Leal de Souza, Pedro H. Presta Dia, Vagner
14 Fonseca, José Lourenço, Luiz Carlos Junior Alcantara, Felicidade Pereira*, Arabela Leal

15
16 **Affiliations:** Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas,
17 Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil (S. Tosta, V.
18 Fonseca, M. Giovanetti, L. Alcantara); Laboratório de Flavivírus, Instituto Oswaldo Cruz,
19 Fundação Oswaldo Cruz, Rio de Janeiro, Brazil (M. Giovanetti, L. Alcantara); Coordenação
20 Geral dos Laboratórios de Saúde Pública/Secretaria de Vigilância em Saúde, Ministério da
21 Saúde, (CGLAB/SVS-MS) Brasília, Distrito Federal, Brazil (V. Fonseca); KwaZulu-Natal
22 Research Innovation and Sequencing Platform (KRISP), School of Laboratory Medicine and
23 Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South

24 Africa Department of Zoology (V. Fonseca); Centro e Informações Estratégicas de Vigilância
25 em Saúde do estado da Bahia (CIEVS) (P. Dias), Secretaria Municipal de Saúde de Salvador (C.
26 Cardoso); Diretoria de Vigilância Epidemiológica do Estado da Bahia (DIVEP) (M. Souza);
27 Secretaria Municipal de Saúde de Irecê (T. Silva); University of Oxford, Oxford OX1 3PS, UK
28 (J. Lourenço); Laboratório Central de Saúde Pública da Bahia – LACEN-BA, Salvador, Bahia,
29 Brazil (S. Tosta, V. Nardy, L. Silva, , M. Gómez, J. Lima, F. Pereira, A. Leal).

30

31 *These authors contributed equally to this article.

32

33 **Abstract**

34 Tracking the spread of SARS-CoV-2 variants of concern is crucial to inform public health efforts
35 and control the ongoing pandemic. Here, we report genetic evidence for circulation of the P.1
36 variant in Northeast Brazil. We advocate for increased active surveillance to ensure adequate
37 control of this variant throughout the country.

38

39 **Text**

40 Since the emergence of the severe acute respiratory syndrome coronavirus-2 (SARS-
41 CoV-2) in 2019, the combination between the unprecedented number of cases and more than
42 500k genomes generated has allowed the identification of hundreds of circulating genetic
43 variants during the pandemic (*1*). Currently, three variants (B.1.1.7 or VOC202012/01, B.1.351
44 or 20H/501Y.V2 and P.1) carrying several mutations in the receptor-binding domain (RBD) of
45 the spike (S) protein, raise concerns about their potential to shift the dynamics and public health
46 impact of the pandemic (2-5). They appear potentially associated with (i) increased

47 transmissibility, (ii) propensity for re-infection, (iii) escape from neutralizing antibodies, and (iv)
48 increased affinity for the human ACE2 receptor (6-8).

49 First identified in January 2021 in travelers from the Amazonas state (North of Brazil)
50 who arrived in Japan, the P.1 variant (alias of B.1.1.28.1) (9), harbors a constellation of 17
51 unique mutations, including three in the receptor binding domain of the spike protein (K417T,
52 E484K, and N501Y). It thus immediately raised concerns to public health authorities over the
53 risk of its unknown potential of faster spreading and/or worsening of coronavirus disease
54 (COVID-19) clinical outcomes. In view of its rapid spread in Brazil and elsewhere (4,5,10) the
55 public health authorities of the Bahia state (Northeast Brazil) conducted an active monitoring for
56 a rapid detection of this variant in the state.

57 Here, we report genetic evidence of the circulation of the P.1 variant in Bahia, by
58 generating 11 SARS-CoV-2 complete genomes from travelers returning from the Amazonas
59 state (North of Brazil). In mid-January 2021, routine genomic surveillance in the Central
60 Laboratory of Health of the Bahia state, started a massive screening of COVID-19 patients and
61 their contacts reporting a travel history to/from the Amazonas state, resulting in eleven suspected
62 SARS-CoV-2, P.1 infections. Viral RNA was extracted from nasopharyngeal swabs and tested
63 for SARS-CoV-2 by multiplex real-time PCR using SARS-CoV-2 assay Allplex™ 2019-nCoV
64 Assay (Seegene) and Charité: SARS-CoV2 (E/RP) (Biomanguinhos). Cycle threshold values
65 (Cts) for the common target between the two assays, gene E, ranged from ≈ 16 to ≈ 35 . We next
66 subjected the qRT-PCR–positive samples to viral genomic amplification and sequencing using
67 the Ion GeneStudio™ S5 Plus Ion Torrent (Life Technologies, USA), according to the
68 manufacturer’s instructions. Consensus sequences from the eleven positive tested samples were
69 generated by *de novo* assembling using Genome Detective (<https://www.genomedetective.com/>)

70 (11). A total of 13,234,814 mapped reads were obtained, resulting in a sequencing mean depth of
71 4824 and a coverage of 98%. The new whole genome sequences generated were assigned,
72 according to the Pangolin (1) lineage classification, as the recently identified P.1 variant (alias of
73 B.1.1.28.1). All patients had recently record travel from the city of Manaus in Amazonas back
74 into the Bahia state. Four patients were from the same family and the other others had no other
75 known connection between them. The data from the eleven patients are described in (**Table 1**).

76

77 We explored the genetic relationship of the newly sequenced P.1 genomes to those of
78 other isolates by phylogenetic inference. To do so, we combined the eleven new isolates
79 (Accession numbers EPI_ISL_1067728 - EPI_ISL_1067738) with all Brazilian SARS-CoV-2
80 (n=1663) genomes, including recently released P.1 genomes (**Figure 1 panels A-B**) (4,5)
81 available on GISAID (<https://www.gisaid.org/>) up to February 21st, 2021. Only genomes
82 >29,000bp and <1% of ambiguities were considered (n=1663). Sequences were aligned using
83 MAFFT (12) and submitted to IQ-TREE for maximum likelihood (ML) phylogenetic analysis
84 (13).

85 Phylogenetic estimation strongly supported placement of the isolates from Bahia within
86 the Brazilian P.1 clade (**Figure 1, panel A**) (SH-aLTR = 1.0). Our results further highlighted
87 that most of our new isolates formed two well supported monophyletic clusters (**Figure 1 panel**
88 **C**). One of those clusters (Cluster i), included samples from the capital city of Salvador, isolated
89 from individuals from the same family of the eleven patients described here (BA49, 50, 51 and
90 52) plus the sample BA57 (SH-aLTR = 1.0). And a second cluster (Cluster ii) which included
91 samples BA53, 54 and 55 that were isolated from patients from the municipalities of João
92 Dourado and Irecê (SH-aLTR = 1.0). Moreover, samples BA56 and BA58 and BA59 appeared to

93 be interspersed among P.1 strains from the North and Southeast Brazil. By combining all the P.1
94 strains already available from distinct Brazilian regions (**Figure 1 panel B, C**) our analysis
95 further revealed that this variant was already detected in the majority of Brazilian regions
96 (starting from the North to Midwest, Northeast, Southeast and South Brazil) highlighting the
97 high connectivity of the country and reinforcing the need of active monitoring to follow the local
98 real-time spread of this new variant of international concern. Finally, among the eleven new
99 genomes, we also identified lineage-specific mutations, including the ones of international
100 concern among the RDB domain: K417T (22812A>C), E484K (23012G>A) and the N501Y
101 (23063A>T) (**Figure 1, panel D**).

102 This report describes the early detection of the SARS-CoV-2 P.1 variant in the Northeast
103 of Brazil (Bahia state) and provides evidence regarding the P.1 circulation across all Brazilian
104 macro-regions that would have occurred within the past two months. The P.1 variant is now
105 known to have emerged in the Amazonas state, but was first detected in travelers arriving in
106 Japan from Brazil. The latter is a testimony for the currently scarce genomic surveillance in
107 Brazil, which has failed to detect the variant before it became a public health emergency in the
108 Amazonas state. Moreover, the eleven individuals here described to carry the P.1 variant in the
109 Bahia state were only detected after a unique active screening initiative that focused on travelers
110 and their recent contacts. This case study should serve as an example of the effectiveness of
111 active surveillance to monitor the importation of genetic variants of importance for public health.
112 Since such initiatives are not only critical to monitor the ongoing spread of known variants but
113 are also necessary to detect the possible emergence of new ones in the near future, continual and
114 revamped investment is needed and will be necessary for adequate public health policy in Brazil.

115

116

117 **Acknowledgments**

118 This work was financed by Laboratório Central de Saúde Pública da Bahia (LACEN-BA) and
119 Secretaria da Saúde do Estado da Bahia (SESAB) and supported by Pan American Health
120 Organization PAHO/WHO. We also thank Centro e Informações Estratégicas de Vigilância em
121 Saúde (CIEVS) in the municipality of Salvador, Bahia. ST is supported by the Coordenação de
122 Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. MG is
123 supported by Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ). JL is
124 supported by a lectureship from the Department of Zoology, University of Oxford. We also
125 would like to thank all the authors who have kindly deposited and shared genome data on
126 GISAID. A table with genome sequence acknowledgments can be found in **Supplementary**
127 **Table 1.**

128

129 **Author Bio**

130 MSc. Fraga is a PhD student at the Federal University of Minas Gerais, Southeast Brazil. Her
131 research focuses on the implementation of next generation sequencing techniques to track the
132 spread of emerging and re-emerging viral pathogens.

133

134 **Ethical Statement**

135 This research was approved by the Ethics Review Committee of the Federal University of Minas
136 Gerais (CEP/CAAE: 32912820.6.1001.5149 approval number).

137

138 **Data Sharing**

139 Newly generated SARS-CoV-2 sequences have been deposited in GISAID under accession
140 numbers EPI_ISL_1067728 to EPI_ISL_1067738.

141 **Declaration of interests**

142 The authors declare no competing interests.

143

144 **AUTHOR CONTRIBUTIONS**

145 **Conception and design:** SFT, MG, FPM, and AL; **Data collection:** SFT, MG, FPM, CC, TS,
146 PD, MS; **Investigations:** ST, LS, VN, MKAG, JGL, CC, MG, VF, JL; **Data Analysis:** ST, MG,
147 VF; **Draft Preparation:** ST, MG, JL, FPM, and AL; **Revision:** LCJA, SFT, MG, FPM, CC, TS,
148 PD, MS.

149

150 **References**

- 151 1. O'Toole Á, Scher E, Anthony Underwood, Ben Jackson, Verity Hill, JT McCrone, Chris
152 Ruis, Khali Abu-Dahab, Ben Taylor, Corin Yeats, Louis du Plessis, David Aanensen,
153 Eddie Holmes, Oliver Pybus, Andrew Rambaut. Pangolin: lineage assignment in an
154 emerging pandemic as an epidemiological tool. (github.com/cov-lineages/pangolin)
- 155 2. Rambaut A, Loman N, Pybus O, et al. Preliminary genomic characterisation of an
156 emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations.
157 [https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-](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)
158 [lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563](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). (accessed Dec 21,
159 2020).
- 160 3. Tegally H, Wilkinson E, Giovanetti M, et al. Emergence and rapid spread of a new severe
161 acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple

- 162 spike mutations in South Africa. medRxiv 2020.12.21.20248640 (Preprint); doi:
163 <https://doi.org/10.1101/2020.12.21.20248640>.
- 164 4. Faria NR, Claro IM, Candido D, et al. Genomic characterisation of an emergent SARS-
165 CoV-2 lineage in Manaus: preliminary findings. *Virological*. Date accessed: February 24,
166 2021. [https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-](https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586)
167 [lineage-in-manaus-preliminary-findings/586](https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586)
- 168 5. Rezende, Paola. “Phylogenetic Relationship of SARS-CoV-2 Sequences from Amazonas
169 with Emerging Brazilian Variants Harboring Mutations E484K and N501Y in the Spike
170 Protein.” *Virological*, 11 Jan. 2021, [virological.org/t/phylogenetic-relationship-of-sars-](https://virological.org/t/phylogenetic-relationship-of-sars-cov-2-sequences-from-amazonas-with-emerging-brazilian-variants-harboring-mutations-e484k-and-n501y-in-the-spike-protein/585)
171 [cov-2-sequences-from-amazonas-with-emerging-brazilian-variants-harboring-mutations-](https://virological.org/t/phylogenetic-relationship-of-sars-cov-2-sequences-from-amazonas-with-emerging-brazilian-variants-harboring-mutations-e484k-and-n501y-in-the-spike-protein/585)
172 [e484k-and-n501y-in-the-spike-protein/585](https://virological.org/t/phylogenetic-relationship-of-sars-cov-2-sequences-from-amazonas-with-emerging-brazilian-variants-harboring-mutations-e484k-and-n501y-in-the-spike-protein/585).
- 173 6. “Emerging SARS-CoV-2 Variants.” Centers for Disease Control and Prevention, Centers
174 for Disease Control and Prevention, [www.cdc.gov/coronavirus/2019-ncov/more/science-](https://www.cdc.gov/coronavirus/2019-ncov/more/science-and-research/scientific-brief-emerging-variants.html)
175 [and-research/scientific-brief-emerging-variants.html](https://www.cdc.gov/coronavirus/2019-ncov/more/science-and-research/scientific-brief-emerging-variants.html).
- 176 7. Vasques Nonaka CK, Franco MM, Gräf T, de Lorenzo Barcia CA, de Ávila Mendonça
177 RN, de Sousa KAF, et al. Genomic evidence of SARS-CoV-2 reinfection involving
178 E484K spike mutation, Brazil. *Emerg Infect Dis*. 2021 May
- 179 8. Weisblum Y, Schmidt F, Zhang F, et al. Escape from neutralizing antibodies by SARS-
180 CoV-2 spike protein variants. *Elife* 2020; 9:3-10.
- 181 9. National Institute of Infectious Diseases, Japan. (2021, January 12.) Brief report: New
182 Variant Strain of SARS-CoV-2 Identified in Travelers from Brazil [Press release].
183 Retrieved from [https://www.niid.go.jp/niid/en/2019-ncov-e/10108-covid19-33-](https://www.niid.go.jp/niid/en/2019-ncov-e/10108-covid19-33-en.htmlexternal)
184 [en.htmlexternal](https://www.niid.go.jp/niid/en/2019-ncov-e/10108-covid19-33-en.htmlexternal)

- 185 10. Pangolin: lineage assignment in an emerging pandemic as an epidemiological tool
186 (https://cov-lineages.org/global_report_P.1.html)
- 187 11. Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, K., ... & de
188 Oliveira, T. (2019). Genome Detective: an automated system for virus identification from
189 high-throughput sequencing data. *Bioinformatics*, 35(5), 871-873.
- 190 12. Nakamura, T., Yamada, K. D., Tomii, K. & Katoh, K. Parallelization of MAFFT for
191 large-scale multiple sequence alignments. *Bioinformatics*. 2018; 34:2490–2492.
- 192 13. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A Fast and
193 Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol*
194 *Biol Evol*. 2015; 32:268–274.

195

196 **Address for correspondence:** Luiz Carlos Junior Alcantara, PhD, Oswaldo Cruz
197 Foundation, luiz.alcantara@ioc.fiocruz.br; Marta Giovanetti, PhD, Oswaldo Cruz
198 Foundation, giovanetti.marta@gmail.com; Felicidade Mota Pereira, Central Laboratory of
199 Health of Bahia State felicidade.pereira@saude.ba.gov.br; Arabela Leal, Central Laboratory
200 of Health of Bahia State, arabelaleal@gmail.com .

201

202 **Figure 1. Genomic detection of the SARS-CoV-2 P.1 variant of concern in Bahia state,**
203 **Northeast Brazil.** a) Maximum likelihood (ML) phylogenetic tree including the 11 new isolates
204 obtained in this study plus n=1663 SARS-CoV-2 Brazilian strains collected up to February 21st,
205 2021. b) Map of Brazil showing the number of P.1 SARS-CoV-2 sequences available by region
206 and state. North region (blue): AM=Amazonas state, RR=Roraima state and RO= Rondônia
207 state; Midwest region (dark green): GO=Goiás state; Southeast region (lilac): SP=São Paulo and
208 South region: SC (light green): Santa Catarina state and RS=Rio Grande do Sul state and
209 Northeast region (red): BA=Bahia state. c) Representation of the zoom of the Brazilian P.1

210 clade. Branch support (SH-aLTR >0.8) is shown at key nodes. d) Variant maps of the P.1
211 lineage-defining-mutations were mapped against the SARS-CoV-2 genome structure. Most
212 common mutations defined as mutations present in more than 90% of the genomes in that group.
213 Mutations of international concern: i) K417T (22812A>C); ii) E484K (23012G>A) and iii)
214 N501Y (23063A>T), among the RDB domain are highlighted in red.

215

216

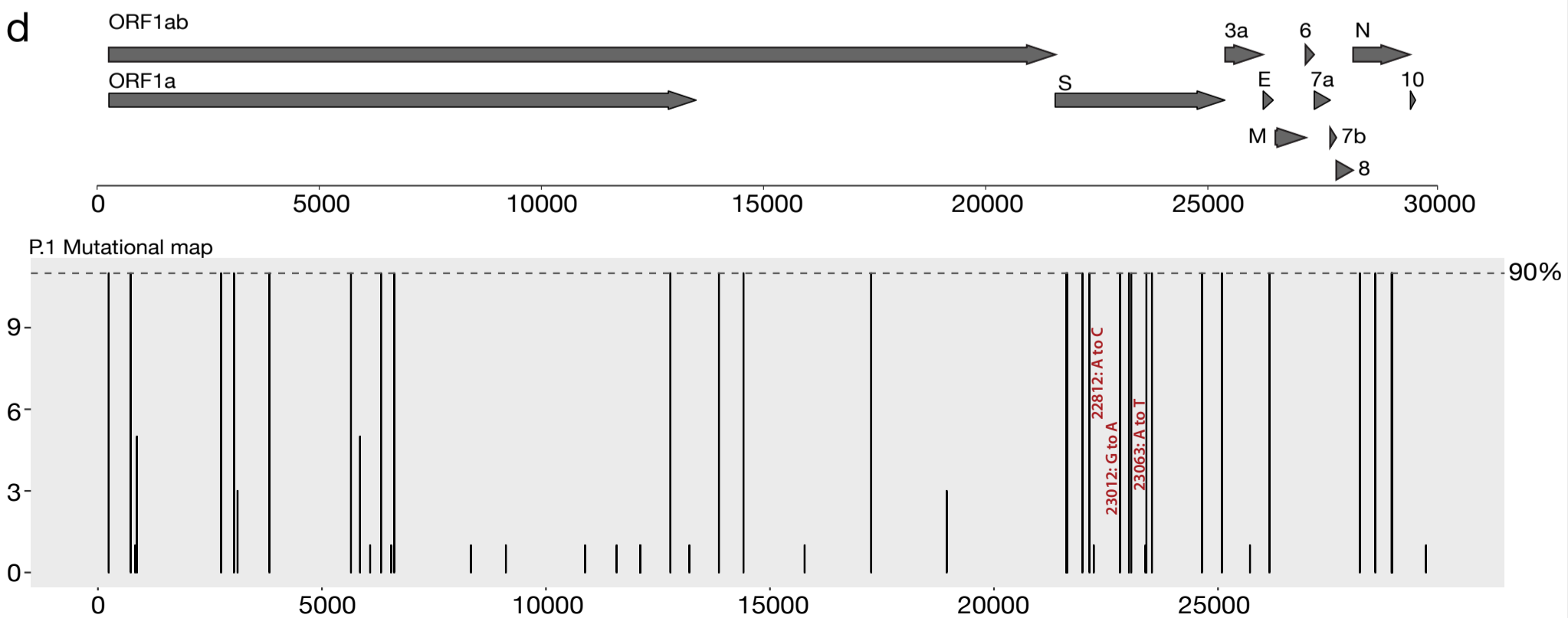
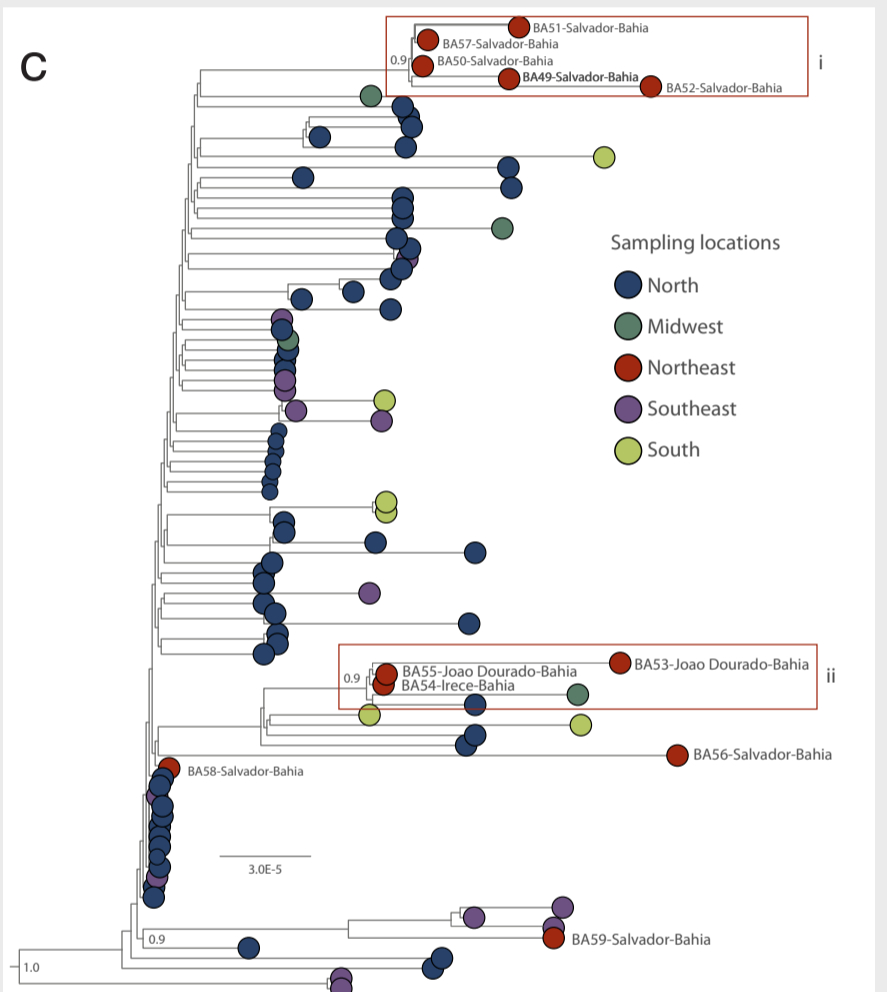
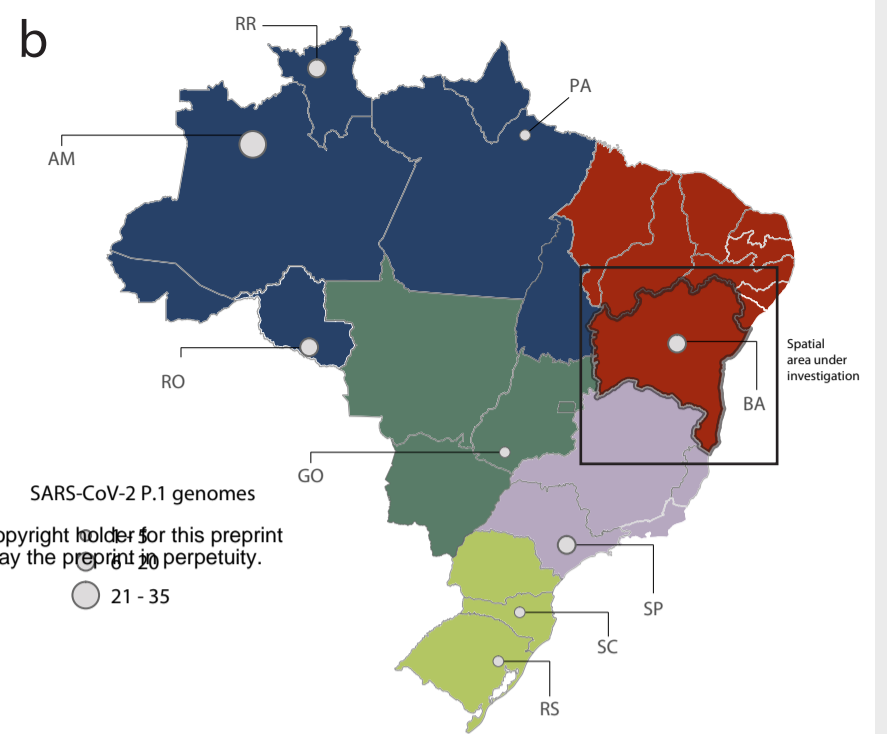


Table 1. Information from patients with travel history from Manaus, Amazonas to state of Bahia.

GISAID ID	ID	Requisition number date	Sample number	First day of symptoms	Colection Date	Sex	Age	Municipality	Cycle threshold (ct)	Reads	Coverage	Depth of Coverage	Identity NT	Identity AA
EPI_ISL_1067737	BA49 ¹	210293001346	290601061	none	26/01/21	M	7	Salvador	CI -16.858 E - 18.197 N - 20.885 RdRp -26.591	1659219	99.8	10179.3	100	99.9
EPI_ISL_1067730	BA50 ¹	210293001349	290601087	11/01/21	26/01/21	F	31	Salvador	CI -28.754 E - 25.728 N - 25.672 RdRp - 28.469	594545	99.8	3706.8	99.9	99.9
EPI_ISL_1067738	BA51 ¹	210173000145	290600024	11/01/21	26/01/21	M	66	Salvador	CI -27.016 E - 24.909 N - 23.841 RdRp - 25.642	412507	82.2	2869	89.9	88.7
EPI_ISL_1067731	BA52 ¹	210293001345	290601051	12/01/21	26/01/21	F	37	Salvador	CI - 27.505 E - 24.558 N - 25.565 RdRp -26.856	1047634	99.2	6779	99.9	99.8
EPI_ISL_1067735	BA 56 ¹	210174000203	290590503	none	16/01/21	M	37	Salvador	CI - 25.285 E - 24.953 N - 25.789 RdRp - 26.624	862894	99.8	5438.3	100	99.9
EPI_ISL_1067729	BA 54	212121000810	290591002	12/01/21	23/01/21	F	56	Irecê	CI -15.407 E -16.941 N - 16.808 RdRp -27.278	625085	99.8	3982.5	100	99.9
EPI_ISL_1067734	BA 55	212114000043	290591058	09/01/21	23/01/21	M	52	João Dourado	CI - 27.088 E - 22.326 N - 23.723 RdRp - 24.493	1848679	99.8	11278.1	100	99.9
EPI_ISL_1067733	BA 53	212114000042	290591052	23/01/21	23/01/21	M	19	João Dourado	E - 22.975 RP - 26.680	1131609	99.8	6979.9	99.9	99.9
EPI_ISL_1067736	BA57	None	SRH-01	None	23/01/21	M	42	Salvador	CI - 39.527 E -17.289 N - 18.974 RdRp -19.226	1416051	99.8	8970.8	99.9	99.9
EPI_ISL_1067728	BA58	None	SRH-02	None	08/01/21	F	53	Salvador	CI - 27.812 E - 29.098 N - 29.766 RdRp - 30.580	2343392	98	14635.7	99.8	99.6
EPI_ISL_1067732	BA59	None	SRH-03	None	18/01/21	M	57	Salvador	CI - 26.759 E - 35.728 N - 32.585 RdRp - 33.041	1293199	99.8	8105.2	100	99.9

1 - Patients from the same family