

1 **Genomic epidemiology reveals the impact of national and international restrictions**  
2 **measures on the SARS-CoV-2 epidemic in Brazil**

3

4 Marta Giovanetti<sup>1,2,\*</sup>&, Svetoslav Nanev Slavov<sup>3,4\*</sup>, Vagner Fonseca<sup>1,5,6,7\*</sup>, Eduan Wilkinson<sup>6,7\*</sup>,  
5 Houriiyah Tegally<sup>6,7\*</sup>, José Salvatore Leister Patané<sup>4\*</sup>, Vincent Louis Viala<sup>4</sup>, James Emmanuel San<sup>6,7</sup>,  
6 Evandra Strazza Rodrigues<sup>3</sup>, Elaine Vieira Santos<sup>3</sup>, Flavia Aburjaile<sup>2,8</sup>, Joilson Xavier<sup>1,8</sup>, Hegger  
7 Fritsch<sup>1,8</sup>, Talita Emile Ribeiro Adelino<sup>1,8</sup>, Felicidade Pereira<sup>9</sup>, Arabela Leal<sup>9</sup>, Felipe Campos de Melo  
8 Iani<sup>8</sup>, Glauco de Carvalho Pereira<sup>8</sup>, Cynthia Vazquez<sup>10</sup>, Gladys Mercedes Estigarribia Sanabria<sup>11,12,13</sup>,  
9 Elaine Cristina de Oliveira<sup>14</sup>, Luiz Demarchi<sup>15</sup>, Julio Croda<sup>16</sup>, Rafael dos Santos Bezerra<sup>3</sup>, Loyze Paola  
10 Oliveira de Lima<sup>4</sup>, Antonio Jorge Martins<sup>4</sup>, Claudia Renata dos Santos Barros<sup>4</sup>, Elaine Cristina  
11 Marqueze<sup>4</sup>, Jardelina de Souza Todao Bernardino<sup>4</sup>, Debora Botequio Moretti<sup>4</sup>, Ricardo Augusto  
12 Brassaloti<sup>17</sup>, Raquel de Lello Rocha Campos Cassano<sup>17</sup>, Pilar Drummond Sampaio Corrêa Mariani<sup>18</sup>, João  
13 Paulo Kitajima<sup>19</sup>, Bibiana Santos<sup>19</sup>, Rodrigo Proto-Siqueira<sup>20</sup>, Vlademir Vicente Cantarelli<sup>21</sup>, Stephane  
14 Tosta<sup>2,9</sup>, Vanessa Brandão Nardy<sup>9</sup>, Luciana Reboredo de Oliveira da Silva<sup>9</sup>, Marcela Kelly Astete  
15 Gómez<sup>9</sup>, Jaqueline Gomes Lima<sup>9</sup>, Adriana Aparecida Ribeiro<sup>8</sup>, Natália Rocha Guimarães<sup>8</sup>, Luiz Takao  
16 Watanabe<sup>14</sup>, Luana Barbosa Da Silva<sup>14</sup>, Raquel da Silva Ferreira<sup>14</sup>, Mara Patricia F. da Penha<sup>22</sup>, María  
17 José Ortega<sup>10</sup>, Andrea Gómez de la Fuente<sup>10</sup>, Shirley Villalba<sup>10</sup>, Juan Torales<sup>10</sup>, María Liz Gamarra<sup>10</sup>,  
18 Carolina Aquino<sup>10</sup>, Gloria Patricia Martínez Figueredo<sup>11,12,13</sup>, Wellington Santos Fava<sup>16</sup>, Ana Rita C.  
19 Motta-Castro<sup>16</sup>, James Venturini<sup>16</sup>, Sandra Maria do Vale Leone de Oliveira<sup>16</sup>, Crhistine Cavalheiro  
20 Maymone Gonçalves<sup>23</sup>, Maria do Carmo Debur Rossa<sup>24</sup>, Guilherme Nardi Becker<sup>24</sup>, Mayra Marinho  
21 Presibella<sup>24</sup>, Nelson Quallio Marques<sup>24</sup>, Irina Nastassja Riediger<sup>24</sup>, Sonia Raboni<sup>25</sup>, Gabriela Mattoso  
22 Coelho<sup>26</sup>, Allan Henrique Depieri Cataneo<sup>26</sup>, Camila Zanluca<sup>26</sup>, Claudia N Duarte dos Santos<sup>26</sup>, Patricia  
23 Akemi Assato<sup>27</sup>, Felipe Allan da Silva da Costa<sup>27</sup>, Mirele Daiana Poleti<sup>28</sup>, Jessika Cristina Chagas  
24 Lesbon<sup>28</sup>, Elisangela Chicaroni Mattos<sup>28</sup>, Cecilia Artico Banho<sup>29</sup>, Lívia Sacchetto<sup>29</sup>, Marília Mazzi  
25 Moraes<sup>29</sup>, Rejane Maria Tommasini Grotto<sup>27,30</sup>, Jayme A. Souza-Neto<sup>27</sup>, Maurício Lacerda Nogueira<sup>29</sup>,

26 Heidge Fukumasu<sup>28</sup>, Luiz Lehmann Coutinho<sup>17</sup>, Rodrigo Tocantins Calado<sup>3</sup>, Raul Machado Neto<sup>4</sup>, Ana  
27 Maria Bispo de Filippis<sup>1</sup>, Rivaldo Venancio da Cunha<sup>31</sup>, Carla Freitas<sup>5</sup>, Cassio Roberto Leonel Peterka<sup>32</sup>,  
28 Cássia de Fátima Rangel Fernandes<sup>33</sup>, Wildo Navegantes de Araújo<sup>34</sup>, Rodrigo Fabiano do Carmo Said<sup>34</sup>,  
29 Maria Almiron<sup>34</sup>, Carlos Frederico Campelo de Albuquerque e Melo<sup>34</sup>, José Lourenço<sup>35,36</sup>, Tulio de  
30 Oliveira<sup>6,7,37,38</sup>, Edward C. Holmes<sup>39</sup>, Ricardo Haddad<sup>4</sup>, Sandra Coccuzzo Sampaio<sup>4^&</sup>, Maria Carolina  
31 Elias<sup>4^&</sup>, Simone Kashima<sup>3^&</sup>, Luiz Carlos Junior de Alcantara<sup>1,2^&</sup>, Dimas Tadeu Covas<sup>3,4^&</sup>.

32

33 \*These authors jointly contributed to this work

34 ^These authors jointly supervised this work

35 Corresponding Authors<sup>&</sup>: [sandra.coccuzzo@butantan.gov.br](mailto:sandra.coccuzzo@butantan.gov.br), [carolina.eliassabbaga@butantan.gov.br](mailto:carolina.eliassabbaga@butantan.gov.br),

36 [skashima@hemocentro.fmrp.usp.br](mailto:skashima@hemocentro.fmrp.usp.br), [luiz.alcantara@ioc.fiocruz.br](mailto:luiz.alcantara@ioc.fiocruz.br), [dimas.covas@butantan.gov.br](mailto:dimas.covas@butantan.gov.br);

37 [giovanetti.marta@gmail.com](mailto:giovanetti.marta@gmail.com)

38

39 <sup>1</sup> Laboratório de Flavivírus, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil; <sup>2</sup>Laboratório de Genética Celular e  
40 Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais,  
41 Brazil; <sup>3</sup>University of São Paulo, Ribeirão Preto Medical School, Blood Center of Ribeirão Preto, Ribeirão Preto,  
42 SP, Brazil; <sup>4</sup>Butantan Institute, São Paulo, Brazil; <sup>5</sup>Coordenação Geral de Laboratórios de Saúde Pública/Secretaria  
43 de Vigilância em Saúde, Ministério da Saúde (CGLAB/SVS-MS) Brasília, Distrito Federal, Brazil; <sup>6</sup>KwaZulu-Natal  
44 Research Innovation and Sequencing Platform (KRISP), School of Laboratory Medicine and Medical Sciences,  
45 College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa; <sup>7</sup>Centre for Epidemic Response  
46 and Innovation (CERI), School of Data Science and Computational Thinking, Stellenbosch University;  
47 Stellenbosch, South Africa; <sup>8</sup>Laboratório Central de Saúde Pública do Estado de Minas Gerais, Fundação Ezequiel  
48 Dias, Belo Horizonte, Minas Gerais, Brazil; <sup>9</sup>Laboratorio Central de Saude Publica da Bahia-LACEN-BA,  
49 Salvador, Bahia, Brazil; <sup>10</sup>Laboratorio Central de Salud Pública, Asunción, Paraguay; <sup>11</sup>Universidad Nacional del  
50 Caaguazú, Instituto Regional de Investigación en Salud; <sup>12</sup>Laboratorio de Biología Molecular, Hospital Regional de  
51 Coronel Oviedo; <sup>13</sup>Ministerio de Salud Pública y Bienestar Social; <sup>14</sup>Laboratório Central de Saúde Pública do Estado  
52 de Mato Grosso, Cuiabá, Brazil; <sup>15</sup>Laboratório Central de Saúde Pública do Estado de Mato Grosso do Sul, Campo

53 Grande, Mato Grosso do Sul, Brazil; <sup>16</sup>Universidade Federal de Mato Grosso do Sul; <sup>17</sup>University of São Paulo,  
54 Centro de Genômica Funcional da ESALQ, Piracicaba, SP, Brazil; <sup>18</sup>NGS Soluções Genômicas, Piracicaba, SP,  
55 Brazil; <sup>19</sup>Mendelics Análise Genômica; <sup>20</sup>Instituto de Biologia Molecular, Laboratório Antonello, Rio Grande do  
56 Sul, Brazil; <sup>21</sup>Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Universidade Feevale, Grupo  
57 Exame Laboratórios, Rio Grande do Sul, Brazil; <sup>22</sup>Vigilância em Saúde do Estado de Mato Grosso, Cuiabá, Brazil;  
58 <sup>23</sup>Secretaria de Saúde do Estado do Mato Grosso do Sul; <sup>24</sup>Laboratório Central do Estado do Paraná (Lacen/PR);  
59 <sup>25</sup>Hospital de Clínicas da Universidade Federal do Paraná, Curitiba, PR; <sup>26</sup>Laboratório de Virologia Molecular -  
60 Instituto Carlos Chagas/Fiocruz PR, Curitiba, PR; <sup>27</sup>São Paulo State University (UNESP), School of Agricultural  
61 Sciences, Department of Bioprocesses and Biotechnology, Botucatu, Brazil; <sup>28</sup>Department of Veterinary Medicine,  
62 School of Animal Science and Food Engineering, University of Sao Paulo, Pirassununga, São Paulo, Brazil;  
63 <sup>29</sup>Laboratório de Pesquisas em Virologia, Departamento de Doenças Dermatológicas, Infecciosas e Parasitárias,  
64 Faculdade de Medicina de São José do Rio Preto; <sup>30</sup>Molecular Biology Laboratory, Applied Biotechnology  
65 Laboratory, Clinical Hospital of the Botucatu Medical School, Brazil; <sup>31</sup>Fundação Oswaldo Cruz, Bio-Manguinhos,  
66 Rio de Janeiro, Rio de Janeiro, Brazil; <sup>32</sup>Coordenação Geral das Arboviroses, Secretaria de Vigilância em  
67 Saúde/Ministério da Saúde (CGARB/SVS-MS), Brasília, Distrito Federal, Brazil; <sup>33</sup>Departamento de Imunização e  
68 Doenças Transmissíveis/Secretaria de Vigilância em Saúde, Ministério da Saúde, Brasília, Distrito Federal, Brazil;  
69 <sup>34</sup>Organização Pan-Americana da Saúde/Organização Mundial da Saúde, Brasília, Distrito Federal, Brazil;  
70 <sup>35</sup>Department of Zoology, Peter Medawar Building, University of Oxford, Oxford, UK; <sup>36</sup>Biosystems and Integrative  
71 Sciences Institute, Universidade de Lisboa, Lisboa, Portugal; <sup>37</sup>Centre for the AIDS Programme of Research in  
72 South Africa (CAPRISA), Durban, South Africa; <sup>38</sup>Department of Global Health, University of Washington, Seattle,  
73 WA, USA; <sup>39</sup>Sydney Institute for Infectious Diseases, School of Life and Environmental Sciences and School of  
74 Medical Sciences, University of Sydney, Sydney, NSW, Australia.

75

76

77

78

79

## 80 **Abstract**

81 Brazil has experienced some of the highest numbers of COVID-19 cases and deaths globally and  
82 from May 2021 made Latin America a pandemic epicenter. Although SARS-CoV-2 established  
83 sustained transmission in Brazil early in the pandemic, important gaps remain in our  
84 understanding of virus transmission dynamics at the national scale. Here, we describe the  
85 genomic epidemiology of SARS-CoV-2 using near-full genomes sampled from 27 Brazilian  
86 states and a bordering country - Paraguay. We show that the early stage of the pandemic in  
87 Brazil was characterised by the co-circulation of multiple viral lineages, linked to multiple  
88 importations predominantly from Europe, and subsequently characterized by large local  
89 transmission clusters. As the epidemic progressed under an absence of effective restriction  
90 measures, there was a local emergence and onward international spread of Variants of Concern  
91 (VOC) and Variants Under Monitoring (VUM), including Gamma (P.1) and Zeta (P.2). In  
92 addition, we provide a preliminary genomic overview of the epidemic in Paraguay, showing  
93 evidence of importation from Brazil. These data reinforce the usefulness and need for the  
94 implementation of widespread genomic surveillance in South America as a toolkit for pandemic  
95 monitoring that provides a means to follow the real-time spread of emerging SARS-CoV-2  
96 variants with possible implications for public health and immunization strategies.

97

## 98 **Introduction**

99 At the end of 2019, a novel respiratory pathogen designated the Severe Acute Respiratory  
100 Syndrome Coronavirus 2 (SARS-CoV-2) emerged in the city of Wuhan, China. Since its  
101 identification, the virus has spread rapidly on a global scale causing an unprecedented  
102 Coronavirus disease 2019 (COVID-19) pandemic (declared on March 11<sup>th</sup> 2020) which has

103 overwhelmed many health care systems<sup>1</sup>. By the mid-February 2022, more than 415 million  
104 cases of COVID-19, with more than 5.84 million associated deaths, have been reported  
105 globally<sup>2</sup>. Brazil has become the epicentre of the COVID-19 epidemic in the Americas, with a  
106 total of 21.2 million cases and a death toll exceeding 591,000 reported cases by the end of  
107 September of 2021, making it one of the countries hardest hit by COVID-19<sup>3</sup>. Critically,  
108 however, a lack of genome sequence data from Brazil has limited our ability to fully understand  
109 transmission dynamics at the national scale.

110 To assess how population restriction measures might have played a role in shaping viral  
111 diversity and evolution during the SARS-CoV-2 epidemic in Brazil, we performed a  
112 phylogenetic and phylogeographic analysis of genomic data from 27 Brazilian states and one  
113 neighbouring country – Paraguay – collected up to September 2021. We show that the early  
114 spread of SARS-CoV-2 in Brazil resulted from initial founder events associated with  
115 international travel (mainly from Europe), which then spread through regional mobility pathways  
116 during the recession of the first epidemic wave. From this, Brazil emerged as an exporter of  
117 SARS-CoV-2 variants to other countries. We further describe the emergence and spread of key  
118 viral Variants of Concern (VOCs; for example, Gamma) and Variants Under Monitoring  
119 (VUMs; for example, Zeta) in Brazil, highlighting how their emergence may have contributed to  
120 a more severe second wave in the country.

121

## 122 **Results**

### 123 *COVID-19 transmission dynamics in Brazil*

124 The first confirmed infection of SARS-CoV-2 in Brazil was on February 26<sup>th</sup> 2020 in the State of  
125 São Paulo (SP), in a traveller returning from Italy (**Fig. 1A**). On March 17<sup>th</sup> 2020 the first

126 COVID-19-related death, a 61-year old male, was reported in the same state<sup>4,5</sup>. Four days later,  
127 all Brazilian states reported at least one confirmed case of COVID-19 and the Brazilian Ministry  
128 of Health (BR-MoH) declared an outbreak of large-scale community transmission of the virus<sup>6</sup>.  
129 By April 10<sup>th</sup>, 2020 the virus had already reached remote locations such as the Yanomami  
130 indigenous community located in the state of Roraima, in northern Brazil<sup>6</sup> (**Fig. 1A**).

131 After the World Health Organization (WHO) declared the outbreak of SARS-CoV-2 as a  
132 public health emergency of international concern on January 30<sup>th</sup> 2020 the Brazilian government  
133 introduced restriction measures to mitigate viral spread (**Fig. 1A**)<sup>7</sup>. The primary measure  
134 involved social isolation, followed by the closure of schools, universities, and non-essential  
135 businesses<sup>8</sup>. Additional measures included the mandatory use of personal protective masks<sup>9</sup>, the  
136 cancellation of events expected to attract large numbers of people and tourists, and only services  
137 considered as essential such as markets and pharmacies remained open<sup>8,10</sup>. However, while the  
138 epidemic was growing, restriction measures were progressively eased to mitigate negative  
139 impacts on the economy. Notably, even during periods of restriction, travel between Brazilian  
140 states largely remained possible, facilitating SARS-CoV-2 transmission throughout the  
141 country<sup>11</sup>. The latter was likely linked to the emergence of new and more contagious viral  
142 lineages documented as VOC - Gamma (lineage P.1) - and VUM - Zeta (lineage P.2), which may  
143 also have contributed to a more severe second wave (**Fig. 1B**)<sup>11,13,14,15</sup>.

144 The COVID-19 death toll in Brazil steadily rose from March 2021. It reached a daily  
145 total of 4,250 on April 2021, the highest number of daily fatalities from COVID-19 worldwide  
146 (**Fig. 1B**). Signs of collapse of the health system were reported in numerous cities among  
147 different regions of the country. The situation worsened considerably following the emergence of  
148 multiple VOCs and VUMs in parallel with a slow progress of the vaccination campaign<sup>16</sup>.

149 Vaccination in Brazil began on January 17<sup>th</sup> 2021 when the Instituto Butantan imported the first  
150 6 million doses of CoronaVac (a whole virus inactivated vaccine) in a collaboration with Sinovac  
151 Biotech (**Fig. 1A**)<sup>17,18</sup>. As of February 16<sup>th</sup> 2022 approximately 71.8% of the Brazilian  
152 population have been vaccinated with the first dose of any of the vaccines currently available  
153 (CoronaVac, AstraZeneca, Pfizer and Janssen), and only 22% had been fully vaccinated (i.e. a  
154 single dose of Janssen or both doses of the other vaccines)<sup>19</sup>.

155 By analysing the total number of COVID-19 notified cases to the end of September 2021,  
156 we observed that the Brazilian region with the highest population density (Southeast) also  
157 contained the highest number of the cases registered in the country, with the state of São Paulo  
158 documenting the largest number of cases (n=4,369,410) in that period (**Fig. 1C**). However, when  
159 we consider the incidence rate (number of reported cases/population) by state, we found that the  
160 Midwest, the least populated region in Brazil, had the highest incidence rate, with 13,604.23  
161 cases/100K inhabitants<sup>1</sup>.

162

### 163 *SARS-CoV-2 genomic data*

164 A total of 3,866 near-full genomes sequences from SARS-CoV-2 RT-qPCR positive  
165 samples were obtained as part of this study. SARS-CoV-2 sequencing spanned February 2020 to  
166 June 2021, sampled from 8 of the 27 Brazilian states (São Paulo=3309; Rio Grande do Sul=48;  
167 Paraná=55; Minas Gerais=80; Mato Grosso do Sul=36; Mato Grosso=51; Bahia=224) and one  
168 neighbouring country, Paraguay (n=63). Almost half of the sequences were from Southeast  
169 Brazil, consistent with it reporting most cases nationally (**Fig. 1C**)<sup>6</sup>. The sequenced samples  
170 were collected from 2,023 females and 1,843 males (**Table S1 and S2**) with a median age of  
171 41.72 years (range: 1 to 90 years of age). All tested samples contained sufficient viral genetic

172 material ( $\geq 2\text{ng}/\mu\text{L}$ ) for library preparation. For positive samples, PCR cycle threshold (Ct)  
173 values were on average 19.93 (range: 10.75 to 30). Sequences had a median genome coverage of  
174 95% (range: 80 to 99.99) and average genome coverage was typically higher for samples with  
175 lower Ct values (**Fig. S1**). Epidemiological information and sequencing statistics of the  
176 generated sequences from Brazil and Paraguay are detailed in **Tables S1** and **S2**, respectively.  
177 Sequences were assigned to 39 different PANGO-lineages based on the proposed dynamic  
178 nomenclature for SARS-CoV-2 lineages (**Fig. S1, Table S1 and S2**) and have been submitted to  
179 GISAID following the WHO guidelines (**Tables S1** and **S2**) (Pangolin version 3.1.7, August  
180 2021).

181

#### 182 *Phylogenetic inference and lineage diversity*

183 The rapid spread of SARS-CoV-2, together with the reported circulation of several VOCs  
184 and VUMs in Brazil prompted the intensification of genomic surveillance by the National  
185 Network for Pandemic Alert of SARS-CoV-2 at the end of December 2020. As of June 30<sup>th</sup> 2021  
186 some 17,135 SARS-CoV-2 genomes from Brazil had been deposited in the GISAID database  
187 from all 27 Brazilian states (**Fig. 2A**). The states with the highest number of sequenced genomes  
188 were São Paulo (n=9,600) and Rio de Janeiro (n=2,031). Although genomic surveillance began  
189 as soon as the first confirmed infections were detected in Brazil, by the end of June 2021 there  
190 was still a paucity of genomic data from some states such as Roraima (n=29), Acre (n=29),  
191 Rondônia (n=37), Tocantins (n=27), Piauí (n=19) and the Federal District (n=33) (**Fig. 2A**). Half  
192 of all Brazilian genomes were deposited in early 2021, suggesting intensified surveillance in the  
193 second wave following the detection of Gamma (and other VOCs (e.g. Alpha/B.1.1.7) and  
194 VUMs (e.g. Zeta) throughout the country (**Fig. 2B**).



195 To obtain a better understanding of the dynamics of SARS-CoV-2 spread in Brazil, we  
196 coupled epidemiological data with phylodynamic analysis for a data set comprising 25,288  
197 available globally representative genomes, including the new genomes obtained in this study  
198 (n=3,866) sampled from December 26<sup>th</sup>, 2019 to June 28<sup>th</sup> 2021 (**Fig. 2C, Fig. S2**). A date-  
199 stamped phylogeny of these data indicated that most of the Brazilian sequences were  
200 interspersed with those introduced from several countries (**Fig. 2C, 2D**). This pattern further  
201 indicated that the co-circulation of multiple SARS-CoV-2 lineages over time was linked to  
202 multiple importations followed by large local transmissions concomitant with a high number of  
203 infections (**Fig. 2C, 2D**).

204 Using an ancestral location state reconstruction on the dated phylogeny we were able to  
205 infer the number of viral imports and exports between Brazil and the rest of the world, and  
206 between individual Brazilian regions (hereafter referred as the North, Northeast, Midwest,  
207 Southeast and South regions) (**Fig. 2D-F**). The bulk of imported introductions (estimated to be  
208 114 independent ones) were largely from Europe (**Fig. 2D**), occurring before the implementation  
209 of restriction measures (April 2020) when the epidemic was rapidly progressing (**Fig. 2D, E**).  
210 However, at least 33 introduction events are inferred to have occurred during enforcement of  
211 preventive measures up to August 2020 (**Fig. 2D, E**), and hence before those measures were  
212 loosened. Finally, although Brazil was a major virus importer, there were approximately 10 times  
213 more inferred exportation events out of Brazil than viral introductions into Brazil (**Fig. 2E**).

214 Our estimates of viral movement within Brazil further suggested that the Southeast region  
215 was the largest contributor of viral exchanges to other regions, comprising approximately 40% of  
216 viral movements from one geographical region to another, followed by the North region that  
217 contributed to approximately 25% of all viral movements. Although these estimates are in line

218 with epidemiological data, this observation is also likely to be influenced by these two regions  
219 having the greatest number of sequences available for analysis.

220

221 *Spatiotemporal spreading of Brazilian VOC (Gamma) and VUM (Zeta)*

222 We next focused on the two Brazilian variants that evolved from the B.1.1.28 lineage and  
223 grew into large transmission clusters during the second wave of the epidemic from January 2021  
224 - the VOC (Gamma/P.1) and VUM (Zeta/P.2). To assess the detailed evolution of these lineages  
225 over time we performed a spatiotemporal phylogeographic analysis using a molecular clock  
226 model.

227 The Gamma VOC was first sampled in Brazil in early January 2021<sup>12,20</sup>. It displayed an  
228 unusual number of lineage-defining mutations in the S protein, including three designated that  
229 may impact transmission, immune escape, and virulence - N501Y, E484K and K417T<sup>21-23</sup>. In  
230 line with previous estimates<sup>12,20</sup>, our phylogeographic analysis suggested that the Gamma variant  
231 emerged around November 21<sup>st</sup> 2020 (95% highest posterior density, 12-29 November 2020) in  
232 Manaus (Amazonas state) in Northern Brazil and spread extensively among Brazilian regions  
233 (**Fig. 3A, C**). Our data reveals multiple introductions of this lineage from the Amazonas state to  
234 Brazil's southeastern, northeastern and midwestern states (**Fig. 3A, C**). By mid-January 2021,  
235 the southeastern and northern regions had also acted as source populations for the introduction of  
236 this variant into the southern region (**Fig 3A, C**).

237 Zeta (P.2) is defined by the presence of the S:E484K mutation in the receptor binding  
238 domain (RBD) and other lineage-defining mutations outside the S protein<sup>15,24</sup>. Although it was  
239 first described in samples from October 2020 in the state of Rio de Janeiro, our phylogeographic  
240 reconstruction suggests that the variant originated in Paraná state in South Brazil late August

241 2020 (95% highest posterior density, 19 August 2020 to 03 September 2020) (**Fig. 4B**). Since  
242 then, Zeta has spread multiple times to much of the southeastern, northeastern, midwestern and  
243 northern Brazilian regions (**Fig. 4D**). Together, our results further suggest that the transmission  
244 dynamics roughly followed patterns of population density, moving most often between the most  
245 populous localities (**Fig. A, B**).

246 By estimating the pattern of migration flows we also examined the potential role of Brazil  
247 as an exporter of the Gamma and Zeta variants to the rest of the world (**Fig. 3**). While the North  
248 region seeded approximately 47% of all Gamma infections into other regions, consistent with it  
249 being where this lineage originated, there is strong evidence both from phylogeographic analysis  
250 (**Fig. 3A**) and ancestral state reconstruction (**Fig. 3C**) that there was considerable subsequent  
251 transfer of Gamma between all regions. Zeta had a different dispersal pattern to Gamma, with  
252 73% of all Zeta movements originating in the Southeast and South regions, consistent with our  
253 phylogeographic reconstruction that this is the geographic source of this lineage (**Fig. 3**).

254 Our analysis further revealed that Brazil has contributed to the international spread of  
255 both variants with at least 316 and 32 exportation events to the rest of the world detected for  
256 Gamma and Zeta variants, respectively (**Fig. 3E**). Consistent with importations, most exports  
257 were to South America (65%) and Europe (14%), followed by Asia (11%), North America (5%),  
258 Africa (2.5%) and Oceania (2.5%) with an increase between January 2021 and March 2021  
259 coinciding with the second wave of infections in Brazil and some relaxation of international  
260 travel restrictions (**Fig. 3E**). As shown elsewhere, these results demonstrate that under relaxation  
261 of travel restrictions SARS-CoV-2 lineages can spread to a diverse range of international  
262 locations<sup>25-30</sup>.

263

264 *Cross-border SARS-CoV-2 transmission from Brazil to Paraguay*

265 To explore the burden of the Brazilian SARS-CoV-2 pandemic on other South American  
266 countries, we provide a preliminary overview of the SARS-CoV-2 epidemic in Paraguay. The  
267 first COVID-19 confirmed case was documented in Paraguay on March 7<sup>th</sup> 2020 in a 32-year-old  
268 man from San Lorenzo, Central Department. Thirteen days later, the first death and the first case  
269 of community transmission were also confirmed. COVID-19 cases in Paraguay rose sharply in  
270 March (**Fig. 4A**) resulting in 100% occupancy of intensive care beds, prompting the government  
271 to declare a strict quarantine to mitigate the spread of the virus<sup>31,32</sup>. By the end of June 2021 a  
272 total number of 460,000 confirmed cases and 15,000 coronavirus-related deaths had been  
273 reported in Paraguay<sup>31</sup>.

274 The COVID-19 epidemic in Paraguay can generally be characterized by three phases:  
275 phase I starting from March 10<sup>th</sup> 2020 characterized by restriction measures; phase II since May  
276 4<sup>th</sup>, 2020, also called “Intelligent/Smart Quarantine” with a gradual return to work and social  
277 activities; and phase III implemented since October 5<sup>th</sup> 2020, known as the “COVID way of  
278 living”, characterized by the relaxation of the restrictions measures and the reopening of national  
279 borders and resumption of international flights<sup>32</sup>.

280 Since the beginning of the epidemic there has been a paucity of whole genomes  
281 sequences from Paraguay, with only n=165 whole genome sequences available on GISAID by  
282 the end of July 2021, about 0.0003% of known cases. This seriously impacts the ability to  
283 characterize the molecular epidemiology of SARS-CoV-2 at a regional level. In collaboration  
284 with the Pan-American Health Organization and the National Public Health Laboratory of  
285 Asunción in Paraguay we obtained a total of 63 near complete genome sequences sampled  
286 between July 2020 and June 2021, representing ~ 40% of the currently available genomes from

287 this country. The selection of the samples was based on the Ct value ( $\leq 30$ ) and availability of  
288 epidemiological metadata, such as date of sample collection, sex, age and municipality of  
289 residence. Thus, by applying these inclusion criteria only 63 positive samples were considered  
290 suitable for this study. As expected, we observed the co-circulation of multiple SARS-CoV-2  
291 lineages (**Fig. 4B**), linked to multiple importations and subsequently characterized by large  
292 transmission clusters.

293 Importantly, our phylogenetic analysis revealed that most of the SARS-CoV-2 variants  
294 currently circulating in Paraguay, including lineages (B.1.1.28; B.1.1.33; Zeta and Gamma),  
295 originally emerged in Brazil (Fig. 4A and Fig.4B), thus suggesting cross-border transmission  
296 from Brazil to Paraguay (**Fig. 4B, C**). This reinforces the importance of nonpharmaceutical  
297 measures in containing and preventing the spread of viral strains into neighbouring countries.

298 As of July 31<sup>th</sup>, 2021 a total of 78% of available genomic sequences from Paraguay were  
299 linked to infections caused by Brazilian variants, with the Gamma VOC being the most prevalent  
300 lineage in the country. As genome sequencing is not widespread, it is difficult to determine how  
301 widely these variants have spread within Paraguay and to other Latin American countries.  
302 However, the abundant COVID-19 cases in Brazil, a country that shares borders with ten  
303 countries, suggests that this risk is likely to be high.

304

## 305 **Discussion**

306 Genome sequencing combined with epidemiological assessment has been widely used as  
307 a tool to track the spread and evolution of SARS-CoV-2, enabling public health authorities to  
308 tailor their control strategies. In addition, genomic data can reveal key properties of emerging  
309 viruses, including precise epidemiological behaviour that patient data alone cannot capture.

310 However, the utility of genomic data is marred by sparse sampling and geographic inequities<sup>33</sup>.  
311 Brazil is an important case study, but SARS-CoV-2 sequences from Brazil are only available  
312 from a tiny fraction of the number of confirmed cases into the country, curbing their utility in a  
313 public health setting and limiting effective control strategies.

314 To help overcome these limitations, we report genomic data obtained by sequencing  
315 3,866 SARS-CoV-2 infection cases confirmed by RT-qPCR from patients residing in 8 of the 27  
316 Brazilian federal states and the neighbouring country of Paraguay. These were analysed together  
317 with n=13,328 and n=102 (up to 30, June 20<sup>th</sup> 2021) publicly available complete genomes  
318 sequences from both countries, respectively. By combining epidemiological and genomic data,  
319 we show how the interplay between the implementation of restriction measures and sustained  
320 SARS-CoV-2 transmission have shaped the Brazilian epidemic over 20 months, including the  
321 dramatic resurgences in case numbers linked with the emergence of VOCs and VUMs. In  
322 particular, we show that multiple independent importations of SARS-CoV-2, predominantly  
323 from Europe, had occurred in Brazil during the early phase of the epidemic (up to April 2020).  
324 We further detected multiple (n=33) international introductions during periods characterized by  
325 the enforcement of preventive measures, demonstrating that these can be ineffective to prevent  
326 importation, as previously reported in other countries<sup>34</sup>.

327 Our analysis shows that viral importations were present throughout the entirety of the  
328 study period, except between May and September 2020 and further suggest that Brazil played  
329 much more as a viral exporter, resulting in 10 times more inferred exportations events than viral  
330 introduction into the country (**Fig. 2E**). Importantly, this was also a period linked to the  
331 emergence of Brazilian VOCs and VUMs. Although we have not measured passenger traffic to  
332 and from Brazil, the viral migration pattern observed appears to be in line with results obtained

333 from other studies combining viral genetic data with epidemiological and travel data, and  
334 provide evidence that the dispersion of respiratory viruses is dictated by human mobility and  
335 population density<sup>26-30</sup>.

336 We also provide the first preliminary overview of the SARS-CoV-2 epidemic in  
337 Paraguay, revealing a high virus connectivity with Brazil. This spread was likely facilitated by  
338 airplane and road networks that form major transport trails linking Brazil to other parts of South  
339 America. The rapid unhampered spread of such VOCs and VUMs into South American  
340 countries, such as Paraguay, indicates that the land-border controls to curb the international  
341 spread of the virus have been largely ineffective and highlights the inherent challenges in  
342 screening cross-border travellers to contain the spread of the virus within this region. Our data  
343 further suggest that national and international travel restrictions lifted at certain points during the  
344 Brazilian epidemic were likely responsible for both virus introduction from international  
345 localities and within-country transmission, with infected travelers acting as carriers.

346 More broadly, our study highlights the utility of SARS-CoV-2 genome sequencing in  
347 COVID-19 outbreak investigation and the need for more comprehensive country-wide studies of  
348 the epidemiology and spread of emerging viral strains. Since the disproportion on the number of  
349 genome sequences available for each Brazilian state/region, our results further highlight the need  
350 of more public investments to strength the genomic capacity across the country.

351 Additionally, the current co-circulation of VUMs and VOCs in Brazil, together with the  
352 slow vaccine rollout, has important implications for public health in this high populous and  
353 regionally important country<sup>35</sup>. Such epidemiological conditions create the perfect environment  
354 for the continued evolution of SARS-CoV-2, potentially enabling the emergence of novel  
355 variants of altered phenotype.

356

## 357 **Materials and Methods**

### 358 *Ethics statement*

359 This research was approved by the Ethics Review Committee of the Pan American Health  
360 Organization (PAHOERC.0344.01), the Federal University of Minas Gerais (CEP/CAAE:  
361 32912820.6.1001.5149), the University of São Paulo (CEP/FZEA: 4.780.992) and the Blood  
362 Center of Ribeirão Preto (CEP/HCRP-FMRP: 50367721.7.1001.5440), and by the Paraguayan  
363 Ministry of Public Health and Social Welfare (MSPyBS/ S.G. no. 0944/18). The availability of  
364 these samples for research purposes during outbreaks of national concern is allowed under the  
365 terms of the 510/2016 Resolution of the National Ethical Committee for Research – Brazilian  
366 Ministry of Health (CONEP - Comissão Nacional de Ética em Pesquisa, Ministério da Saúde)  
367 that authorizes, without the necessity of an informed consent, the use of clinical samples  
368 collected in the Brazilian Central Public Health Laboratories to accelerate knowledge building  
369 and contribute to surveillance and outbreak response. The samples processed in this study were  
370 obtained anonymously from material exceeding the routine diagnosis in Brazilian public health  
371 laboratories that belong to the public network within BrMoH.

372

### 373 *Genomic surveillance network*

374 As a part of the National Network for Pandemic Alert of SARS-CoV-2, we performed genomic  
375 monitoring to rapidly understand the spread of SARS-CoV-2 at the national and cross border  
376 levels. This was performed in collaboration with the Central Laboratory of Public Health  
377 (LACEN) from the states of Bahia (BA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas  
378 Gerais (MG), as well as Blood Center of Ribeirão Preto, the Fiocruz Paraná (PR), the Butantan



379 Institute from the state of São Paulo (SP) state, the University of São Paulo and the Central  
380 Laboratory of Health of Paraguay.

381

382

383

384 *Sample collection and molecular diagnostic assays*

385 Convenience clinical samples from patients with suspected SARS-CoV-2 infection from patients  
386 residing in 8 of the 27 Brazilian federal states and from Asunción, Paraguay, and collected  
387 between July 2020 and June 2021 were provided for diagnostic and genome sequencing  
388 purposes. Viral RNA was extracted from nasopharyngeal swabs using an automated protocol and  
389 tested for SARS-CoV-2 by multiplex real-time PCR assays: (i) the Allplex 2019-nCoV Assay  
390 (Seegene) targeting the envelope (E), the RNA dependent RNA polymerase (RdRp) and the  
391 nucleocapsid (N) genes; (ii) the Charité: SARS-CoV2 (E/RP) assay (Bio-Manguinhos/Fiocruz)  
392 targeting the E gene, and (iii) the GeneFinder COVID-19 Plus RealAmp Kit (Osang Healthcare,  
393 South Korea) supplied by the BrMoH, Butantan Institute and the Pan-American Health  
394 Organization (OPAS).

395

396 *cDNA synthesis and whole genome sequencing*

397 Samples were selected for sequencing based on the Ct value ( $\leq 30$ ) and availability of  
398 epidemiological metadata, such as date of sample collection, sex, age and municipality of  
399 residence. The preparation of SARS-CoV-2 genomic libraries was performed using both the  
400 Illumina COVIDSeq test following the manufacturer's instructions and nanopore sequencing  
401 using the ARTIC Network primal scheme (<https://github.com/artic-network/artic->

402 ncov2019/tree/master/primer\_schemes/nCoV-2019/V3)<sup>36</sup>. The normalized libraries were loaded  
403 onto a 300-cycle MiSeq Reagent Kit v2 and run on the Illumina MiSeq instrument (Illumina, San  
404 Diego, CA, USA).

405         Due to regional characteristics and the accessibility of resources of the different Brazilian  
406 and Paraguayan laboratories, we also applied field SARS-CoV-2 sequencing using the Oxford  
407 Nanopore MinION technology. In this case the SuperScript IV Reverse Transcriptase kit  
408 (Invitrogen) was initially used for cDNA synthesis following the manufacturer's instructions.  
409 The cDNA generated was subjected to multiplex PCR sequencing using the Q5 High Fidelity  
410 Hot-Start DNA Polymerase (New England Biolabs) and a set of specific primers designed by the  
411 ARTIC Network for sequencing the complete SARS-CoV-2 genome (Artic Network version  
412 3)<sup>36</sup>. PCR conditions have been previously reported in<sup>36</sup>. All experiments were performed in a  
413 biosafety level-2 cabinet. Amplicons were purified using 1x AMPure XP Beads (Beckman  
414 Coulter) and quantified on a Qubit 3.0 fluorimeter (ThermoFisher Scientific) using Qubit™  
415 dsDNA HS Assay Kit (ThermoFisher Scientific). DNA library preparation was performed using  
416 the Ligation Sequencing Kit LSK109(Oxford Nanopore Technologies) and the Native Barcoding  
417 Kit (NBD104 and NBD114, Oxford Nanopore Technologies). Sequencing libraries were loaded  
418 into a R9.4 flow cell (Oxford Nanopore Technologies). In each sequencing run, we used negative  
419 controls to prevent and check for possible contamination with less than 2% mean coverage.

420

#### 421 *Generation of consensus sequences from Illumina and nanopore*

422 The genome assembly pipeline for Illumina reads involved: (i) read trimming and filtering using  
423 Trimmomatic<sup>37</sup>; (ii) minimap2<sup>38</sup> for read mapping against the reference strain (Wuhan-Hu-1  
424 genome reference - NCBI accession NC\_045512.2); (iii) samtools<sup>39</sup> for sorting and indexing;

425 (iv) Pilon<sup>40</sup> for improving the indel detection; (v) bwa mem<sup>41</sup> for remapping against Pilon's  
426 generated consensus; (vi) samtools mpileup to generate alignment quality values; (vii) seqtk<sup>42</sup> to  
427 generate a quasi-final genome version; (viii) bwa mem for a 3rd round of remapping reads  
428 against the quasi-final genome; and (ix) samtools depth to assess position depths given the \*.bam  
429 file from the previous step (nucleotide positions with read depth < 5 are denoted as "N").

430 Oxford Nanopore sequencing raw files were basecalled using Guppy v3.4.5 and barcode  
431 demultiplexing was performed using qcat. Consensus sequences were generated by *de novo*  
432 assembling using Genome Detective<sup>43</sup> that uses DIAMOND to identify and classify candidate  
433 viral reads in broad taxonomic units, using the viral subset of the Swissprot UniRef protein  
434 database. Candidate reads were next assigned to candidate reference sequences using NCBI  
435 blastn and aligned using AGA (Annotated Genome Aligner) and MAFFT. Final contigs and  
436 consensus sequences are then available as FASTA files.

437

#### 438 *Data quality control and global data set collection*

439 To ensure the quality of the genome sequences generated in this study and to guarantee the  
440 highest possible phylogenetic accuracy, only genomes >29,000bp and <1% of ambiguities were  
441 considered (n=3,866). Multiple sequence alignment was performed with MAFFT<sup>44</sup> and a  
442 preliminary phylogenetic tree was inferred using the maximum likelihood method in IQ-TREE2,  
443 employing the GTR+I model of nucleotide substitution. Prior to further phylogenetic analysis  
444 our data set was also assessed for both sequences with low data quality (e.g. with assembling  
445 issues, sequencing and alignment errors, data annotation errors and sample contamination) and  
446 for molecular clock signal (i.e. temporal structure) using TempEst v1.5.3<sup>45</sup>.

447 We appended the 3,866 genome sequences newly generated under this project with an  
448 extensive reference data set of SARS-CoV-2 sequences sampled globally collected since the start  
449 of the outbreak, including n=13,328 near-complete genomes from Brazil and n=102 from  
450 Paraguay (sampled up to June 30, 2021). A unique set of external references (n=7,992) were  
451 obtained by including all the non-Brazilian and non-Paraguayan sequences from the global and  
452 South American NextStrain builds (date of access; June 30<sup>th</sup>, 2021). These global references  
453 were further supplemented by the inclusion of a random sample of P.1 and P.2 lineages from  
454 around the world. This approach was taken to ensure we captured sufficient global sequence  
455 diversity while at the same time enriching our data set with South American sequences with  
456 whom Brazil and Paraguay share borders. Additionally, during our down sampling strategy, we  
457 took into account that the different sequencing efforts employed by different Brazilian states  
458 resulted in differing numbers of available genomes for each state/region. For this reason, we  
459 selected reference sequences considering both the number of reported cases for each state and the  
460 number of available genome sequences, aiming at building a more representative data set (**Fig.**  
461 **S3**).

462

### 463 *Phylogenetic analysis*

464 Sequences were aligned using MAFFT<sup>44</sup> and submitted to IQ-TREE2 for maximum likelihood  
465 (ML) phylogenetic analysis<sup>46</sup> employing the general time reversible (GTR) model of nucleotide  
466 substitution and a proportion of invariable sites (+I) as selected by the ModelFinder application.  
467 Branch support was assessed using the approximate likelihood-ratio test based on the bootstrap  
468 and the Shimodaira–Hasegawa-like procedure (SH-aLRT) with 1,000 replicates.

469 The raw ML tree topology was then used to estimate the number of viral transmission  
470 events between various Brazilian regions and the rest of the world. TreeTime<sup>47</sup> was used to  
471 transform this ML tree topology into a dated tree using a constant mean rate of  $8.0 \times 10^{-4}$   
472 nucleotide substitutions per site per year, after the exclusion of outlier sequences. A migration  
473 model was fitted to the resulting time-scaled phylogenetic tree in TreeTime, mapping country  
474 and regional locations to tips and internal nodes<sup>47</sup>. Using the resulting annotated tree topology  
475 we were able to count the number of transitions (i.e. virus importations and exportations) within  
476 different Brazilian regions and between Brazil and the rest of the world. Importantly, this  
477 analysis was not dependent on a monophyletic clustering of SARS-CoV-2 in Brazil. To provide  
478 a measure of confidence in the time and source of viral transitions, we performed the discrete  
479 ancestral state reconstruction on 10 bootstrap replicate trees (**Fig. S4-S6**).

480

#### 481 *Lineage classification*

482 We used the dynamic lineage classification as specified in the Phylogenetic Assignment of  
483 Named Global Outbreak LINEages (Pangolin version 3.1.7) protocol<sup>48</sup>. This was aimed at  
484 identifying the most epidemiologically important lineages of SARS-CoV-2 circulating within  
485 South America (Brazil and Paraguay). Both VOCs and VUMs were designated based on the  
486 World Health Organization framework as of July, 2021.

487

#### 488 *Phylogeographic reconstruction*

489 VOCs and VUMs that emerged in Brazil (i.e. P.1 and P.2) were identified as monophyletic  
490 groups on the time-resolved phylogenetic trees (**Fig.2 and Fig. S7**). Genome sequences from  
491 these lineages were then extracted to infer continuous phylogeography histories using the

492 Markov chain Monte Carlo (MCMC) method in BEAST v1.10.4<sup>49</sup>, employing the HKY +  $\Gamma$   
493 \_nucleotide substitution model and a strict molecular clock in BEAST v.1.10.4 that assumes constant  
494 evolutionary rates throughout the phylogeny.

495         Due to the huge size of the largest clusters we down-sampled them to <600 taxa per clade  
496 to infer phylogeographies within a manageable time frame. Accordingly, we retained the ten  
497 earliest sequences from each unique sampling location within Brazil and a random distribution  
498 over the remaining samples from each location (n=531 for P.1 and n=428 for P.2). Briefly,  
499 sequences from the subsampled cluster were aligned using MAFFT and preliminary ML trees  
500 were inferred in IQ-TREE2 as described above. Prior to phylogeographic analysis, each lineage  
501 was also assessed for molecular clock signal using the root-to-tip regression method available in  
502 TempEst v1.5.3<sup>45</sup> following the removal of potential outliers that may violate the molecular  
503 clock assumption. We accepted temporal structure when the correlation coefficient was  $> 0.2$ <sup>50</sup>.  
504 Linear regression of root-to-tip genetic distances against sampling dates indicated that the SARS-  
505 CoV-2 sequences from Gamma and Zeta evolved in a relatively clock-like manner ( $r^2= 0.5$ ;  
506 coefficient correlation= 0.45 for Gamma and  $r^2= 0.5$ ; coefficient correlation= 0.40 for Zeta) (**Fig.**  
507 **S7**).

508         We modelled the phylogenetic diffusion and spread of these lineages within Brazil by  
509 analysing localized transmission (between Brazilian regions) using a flexible relaxed random  
510 walk (RRW) diffusion model<sup>51,52</sup> that accommodates branch-specific variation in rates of  
511 dispersal with a Cauchy distribution and a jitter window size of 0.01<sup>52</sup>. The choice of Cauchy  
512 distribution was based on recent studies demonstrating that it is successful for this kind of  
513 analyses based on genomes of SARS-CoV-2 and its variants<sup>4,50,54-56</sup>. For each sequence, latitude  
514 and longitude coordinates were attributed. MCMC analyses were set up in BEAST v1.10.4,  
515 running in duplicate for 100 million iterations and sampling every 10,000 steps in the chain.

516 Convergence for each run was assessed in Tracer v1.7.1 (ESS for all relevant model parameters  
517 >200). Maximum clade credibility trees for each run were summarized using TreeAnnotator after  
518 discarding the initial 10% as burn-in. While the sampling is relatively homogeneous among  
519 sampled locations, the phylogeographic reconstruction will of course remain sampling-  
520 dependent. Hence, differences in sampling effort may have impacted the estimated transition  
521 frequencies between locations. Despite this caveat, the phylogeographic analysis still provides  
522 important information on the history and dynamics of dispersal of the lineages sampled, which  
523 can in turn provide insights into the connectivity of locations along the transmission network.  
524 Finally, we used the R package “seraphim”<sup>55</sup> to extract and map spatiotemporal information  
525 embedded in the posterior trees. Note that a transmission link on the phylogeographic map can  
526 denote one or more transmission events depending on the phylogeographic inference.

527

#### 528 *SARS-CoV-2 time series from Brazil*

529 Data from weekly notified and laboratory confirmed cases of infection by SARS-CoV-2 in  
530 Brazil, were supplied by the Brazilian Ministry of Health, as made available by the COVIDA  
531 network at <https://github.com/wcota/covid19br>. For convenience, the geographical locations  
532 were aggregated by Brazilian macro regions: North, Northeast, Southeast, South, and Midwest.

533

#### 534 **References**

- 535 1. Coronavirus Disease (COVID-19) Situation Reports.  
536 <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
- 537 2. COVID-19 Map. *Johns Hopkins Coronavirus Resource Center*  
538 <https://coronavirus.jhu.edu/map.html>.

- 539 3. COVID-19 Situation Reports - PAHO/WHO | Pan American Health Organization.  
540 <https://www.paho.org/en/covid-19-situation-reports>.
- 541 4. Candido DS, et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science*. **4**,  
542 369(2020)
- 543 5. Xavier J, et al. The ongoing COVID-19 epidemic in Minas Gerais, Brazil: insights from  
544 epidemiological data and SARS-CoV-2 whole genome sequencing. *Emerg Microbes Infect*.  
545 **9**,1824(2020).
- 546 6. Coronavírus Brasil. <https://covid.saude.gov.br/>.
- 547 7. Cota, W. Monitoring the number of COVID-19 cases in Brazil. *Wesley Cota - Physics and*  
548 *Complex Networks* <https://covid19br.wcota.me/en/>.
- 549 8. COVID-19 outbreak in Brazil: adherence to national preventive measures and impact on  
550 people's lives, an online survey | BMC Public Health | Full Text.  
551 <https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-10222-z>.
- 552 9. Pinto Neto, O. *et al.* Mathematical model of COVID-19 intervention scenarios for São  
553 Paulo—Brazil. *Nat Commun* **12**, 418 (2021).
- 554 10. Monteiro de Oliveira, M., Fuller, T. L., Brasil, P., Gabaglia, C. R. & Nielsen-Saines, K.  
555 Controlling the COVID-19 pandemic in Brazil: a challenge of continental proportions. *Nat*  
556 *Med* **26**, 1505–1506 (2020).
- 557 11. Candido DDS. et al. Routes for COVID-19 importation in Brazil. *J Travel Med* **18**, 27  
558 (2020).
- 559 12. Faria NR, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus,  
560 Brazil. *Science* **21**,372 (2021).



- 561 13. Naveca FG, et al. COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic  
562 lineages and P.1 emergence. COVID-19 in Amazonas, Brazil, was driven by the persistence  
563 of endemic lineages and P.1 emergence. *Nat Med*. **27**,1230 (2021).
- 564 14. Sant'Anna FH, et al. Emergence of the novel SARS-CoV-2 lineage VUI-NP13L and massive  
565 spread of P.2 in South Brazil. *Emerg Microbes Infec*. **10**, 1431(2021).
- 566 15. Voloch, C. M. *et al*. Genomic characterization of a novel SARS-CoV-2 lineage from Rio de  
567 Janeiro, Brazil. *J Virol* JVI.00119-21 (2021).
- 568 16. de Oliveira Andrade, R. Covid-19 is causing the collapse of Brazil's national health service.  
569 *BMJ* **370**, m3032 (2020).
- 570 17. Vacinação contra a Covid-19 no Brasil - #PÁTRIAVACINADA.  
571 <https://www.gov.br/saude/pt-br/vacinacao>.
- 572 18. Vacinas contra a COVID-19 - OPAS/OMS | Organização Pan-Americana da Saúde.  
573 <https://www.paho.org/pt/vacinas-contracovid-19>.
- 574 19. Brazil – COVID19 Vaccine Tracker. <https://covid19.trackvaccines.org/country/brazil/>.
- 575 20. Naveca, F. G. et al. COVID-19 in Amazonas, Brazil, was driven by the persistence of  
576 endemic lineages and P.1 emergence. *Nat Med* **27**, 1230 (2021).
- 577 21. Weisblum, Y. et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein  
578 variants. *eLife* **9**, e61312 (2020).
- 579 22. Nonaka, C. K. V. et al. Genomic Evidence of SARS-CoV-2 Reinfection Involving E484K  
580 Spike Mutation, Brazil. *Emerging Infectious Diseases journal*. **5**, 27 (2021).
- 581 23. CDC. Coronavirus Disease 2019 (COVID-19). *Centers for Disease Control and Prevention*  
582 [https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-](https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-emerging-variants.html)  
583 [emerging-variants.html](https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-emerging-variants.html) (2020).

- 584 24. Sant'Anna, F. H. *et al.* Emergence of the novel SARS-CoV-2 lineage VUI-NP13L and  
585 massive spread of P.2 in South Brazil. *Emerging Microbes & Infections* **10**, 1431–1440  
586 (2021).
- 587 25. Kraemer MUG, et al. The effect of human mobility and control measures on the COVID-19  
588 epidemic in China. *Science*. **1**, 368. (2020).
- 589 26. Lemey P, et al. Unifying viral genetics and human transportation data to predict the global  
590 transmission dynamics of human influenza H3N2. *PLoS Pathog.* **5**, 29 (2014).
- 591 27. Lu L, et al. Quantifying predictors for the spatial diffusion of avian influenza virus in China.  
592 *BMC Evol Biol.* **67**, 198. (2017).
- 593 28. Ganasegeran K, et al. Influence of Population Density for COVID-19 Spread in Malaysia:  
594 An Ecological Study. *Int J Environ Res Public Health.* **4**, 76 (2021).
- 595 29. Magee D, et al. Combining phylogeography and spatial epidemiology to uncover predictors  
596 of H5N1 influenza A virus diffusion. *Arch Virol.* **31**, 96 (2015).
- 597 30. du Plessis L, et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the  
598 UK. *Science.* **12**, 371 (2021).
- 599 31. PAHO Paraguay COVID-19 Response. [https://paho-covid19-response-](https://paho-covid19-response-who.hub.arcgis.com/pages/paho-paraguay-covid-19-response)  
600 [who.hub.arcgis.com/pages/paho-paraguay-covid-19-response](https://paho-covid19-response-who.hub.arcgis.com/pages/paho-paraguay-covid-19-response).
- 601 32. Paraguay: WHO Coronavirus Disease (COVID-19) Dashboard With Vaccination Data.  
602 <https://covid19.who.int>.
- 603 33. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data – from  
604 vision to reality. *Euro Surveill* **22**, 30494 (2017).
- 605 34. Jorge DCP, et al. Assessing the nationwide impact of COVID-19 mitigation policies on the  
606 transmission rate of SARS-CoV-2 in Brazil. *Epidemics.* **35**,100.

- 607 35. Victora PC, et al. Estimating the early impact of vaccination against COVID-19 on deaths  
608 among elderly people in Brazil: Analyses of routinely-collected data on vaccine coverage  
609 and mortality. *EClinicalMedicine*. **38**,101.(2021)
- 610 36. Quick, J. nCoV-2019 sequencing protocol. (2020) doi:10.17504/protocols.io.bbmuik6w.
- 611 37. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence  
612 data. *Bioinformatics*. **1**, 30 (2014).
- 613 38. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. **15**,34 (2018).
- 614 39. Danecek P, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 16, 10 (2021).
- 615 40. Walker BJ, et al. Pilon: an integrated tool for comprehensive microbial variant detection and  
616 genome assembly improvement. *PLoS One*. **19**,9 (2014).
- 617 41. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.  
618 *Bioinformatics*. **15**, 25 (2009).
- 619 42. Shen W, et al. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File  
620 Manipulation. *PLoS One*. 11,10 (2016).
- 621 43. Vilsker M, et al. Genome Detective: an automated system for virus identification from high-  
622 throughput sequencing data. *Bioinformatics*. **1**, 35 (2019).
- 623 44. Katoh K, et al. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in  
624 Performance and Usability. *Mol Biol Evol*. **30**, 772 (2013)
- 625 45. Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure  
626 of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* **2**, vew007  
627 (2016).
- 628 46. Minh BQ, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference  
629 in the Genomic Era. *Mol Biol Evol*. **1**, 37 (2020).

- 630 47. Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: Maximum-likelihood phylodynamic  
631 analysis. *Virus Evol* **4**, vex042 (2018).
- 632 48. Rambaut A, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist  
633 genomic epidemiology | *Nat Microbiol.* **5**, 1403 (2020).
- 634 49. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian  
635 phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2**,298 (2018).
- 636 50. Tegally, H., Wilkinson, E., Lessells, R.J. et al. Sixteen novel lineages of SARS-CoV-2 in  
637 South Africa. *Nat Med.* **27**, 440 (2021).
- 638 51. Lemey, P., Rambaut, A., Welch, J. J. & Suchard, M. A. Phylogeography takes a relaxed  
639 random walk in continuous space and time. *Mol Biol Evol* **27**, 1877–1885 (2010).
- 640 52. Dellicour S, et al. Relax, Keep Walking - A Practical Guide to Continuous Phylogeographic  
641 Inference with BEAST. *Mol Biol Evol.* **29**, 38 (2021).
- 642 53. Dellicour, S., Rose, R., Faria, N. R., Lemey, P. & Pybus, O. G. SERAPHIM: studying  
643 environmental rasters and phylogenetically informed movements. *Bioinformatics* **32**, 3204–3206  
644 (2016).
- 645 54. Tegally H, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature*. doi:  
646 10.1038/s41586-021-03402-9;
- 647 55. Viana R, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern  
648 Africa. *Nature*, doi: 10.1038/s41586-022-04411-y.
- 649 56. Dellicour S, Hong SL, Vrancken B, Chaillon A, Gill MS, Maurano MT, Ramaswami S,  
650 Zappile P, Marier C, Harkins GW, Baele G, Duerr R, Heguy A. Dispersal dynamics of SARS-  
651 CoV-2 lineages during the first epidemic wave in New York City. *PLoS Pathog.* 2021  
652 20;17(5):e1009571. doi: 10.1371/journal.ppat.1009571.

653

654 **Acknowledgments:** We thank all those who have kindly deposited and shared genome data on  
655 GISAID (Table S3). The authors acknowledge the technical support of Luciana de Araujo  
656 Pimenta, Luiz Aurelio de Campos Crispin, Gabriela Mauric Frossard Ribeiro, Glaucia Maria  
657 Rodrigues Borges, Mariane Evaristo e Josiane Serrano Borges from the National Network for  
658 Pandemic Alert of SARS-CoV-2 and the contribution of all employees of General Coordination  
659 of Public Health Laboratories and professionals of Public Health Laboratories of Brazil and  
660 Network for Pandemic Alert of Emerging SARS-CoV-2 Variants for their contribution towards  
661 the sequencing effort and for their commitment and work during the fight of the COVID-19  
662 pandemic.

663

664 **Funding:** This study was financed by São Paulo Research Foundation (FAPESP) (Grants  
665 Number: 2020/10127-1 and 2020/05367-3, 2013/08135-2), the Central Public Health  
666 Laboratories, Blood Center of Ribeirão Preto and supported by the Brazilian Ministry of Health  
667 and the Pan American Health Organization PAHO/WHO (APO21-00010098). This work was  
668 supported in part through National Institutes of Health USA grant U01 AI151698 for the United  
669 World Arbovirus Research Network (UWARN), the CRP- ICGEB RESEARCH GRANT 2020  
670 Project CRP/BRA20-03, Contract CRP/20/03, the Oswaldo Cruz Foundation VPGDI-027-FIO-  
671 20-2-2-30, the Brazilian Ministry of Health (SCON2021-00180), the CNPq (426559/2018-5) and  
672 the Faperj E-26/202.930/2016. MG and LCJA are supported by Fundação de Amparo à Pesquisa  
673 do Estado do Rio de Janeiro (FAPERJ) (E-26/202.248/2018(238504);  
674 E26/202.665/2019(247400)). JX is supported by Coordenação de Aperfeiçoamento de Pessoal de  
675 Nível Superior (CAPES) - Finance Code 001. JL is supported by a lectureship by the Department

676 of Zoology of the University of Oxford. VF is supported by Coordenação de Aperfeiçoamento de  
677 Pessoal de Nível Superior (CAPES) (88882.349290/2019-01) and a Flagship grant from the  
678 South African Medical Research Council (MRC-RFA-UFSP-01-2013/UKZN HIVEPI). ECH is  
679 funded by an ARC Australian Laureate Fellowship (FL170100022). MLN is supported by  
680 FAPESP (Grant 20/04836-0) and partially supported by a NIH grant to CreateNeo  
681 (U01AI151807).

682

683

684

685

686 **Author contributions:**

687 Molecular screening and produced SARS-CoV-2 genomic data: MG, SNS, VLV, ESR, EVS,  
688 FA, JX, HF, TERA, LPOL, AJM, CRSB, ECM, JSTB, DBM, RAB, RLRCC, PDSCM, JPK, BS,  
689 RPS, VVC, ST, VBN, LROS, MKAG, JQL, AAR, NRG, LTW, LBS, RSF, MPFP, MJO, SGF,  
690 SV, CA, GPMG, WSF, ARMC, JV, SMVLO, CCMG, MCDR, GNB, MPG, NQM, INR, SR,  
691 GMC, AHDC, CZ, CDS, PAA, FASC, MDP, JCCL, ECM, CAB, LS, MMM, RMTG, JASN,  
692 MLN, HF, LLC, RTC, RMN; Collected samples and curated metadata: CF, CRLP, CFRF, WN,  
693 RFCS, MA, CCAM, RH, SCS, MCE, SK, LCJA, DTC, JASN, HF, LLC, FP, AL, FCMI, GCP,  
694 CV, GMES, ECO, LD, JC; Analysed the data: MG, VF, EW, HT, JSPL, EJS, RB, JL; Helped  
695 with study design and data interpretation: MG, VF, EW, HT, EH, JL, TdO, SCS, MCE, SK,  
696 LCJA, DTC; Wrote the initial manuscript, which was reviewed by all authors: MG, SNS, SCS,  
697 MCE, SK, LCJA, DTC.

698

699 **Competing interests:** The authors declare no competing interests.

700

701 **Data and materials availability:** All sequences that were generated and used in the present  
702 study are listed in table S1-S3 (accessible on the GitHub repository) along with their GISAID  
703 sequence IDs, dates of sampling, the originating and submitting laboratories and main authors.

704 All input files (e.g. alignments or XML files), all resulting output files and scripts used in the  
705 study are shared publicly on GitHub

706 ([https://github.com/genomicsurveillance/Genomic epidemiology reveals how restriction meas-  
707 ures shaped the SARS-CoV-2 epidemic in Brazil](https://github.com/genomicsurveillance/Genomic_epidemiology_reveals_how_restriction_measures_shaped_the_SARS-CoV-2_epidemic_in_Brazil)).

708

#### 709 **ADDITIONAL INFORMATION**

710 **Figure S1. Sequencing statistics and lineage assignment of the SARS-CoV-2 genomes**

711 **generated in this study.** (A) Genome coverage plotted against RT-qPCR cycle threshold value.

712 Colours represent genomes obtained from Brazil (green) and Paraguay (orange), respectively.

713 (B) Number of genomes obtained in this study from Brazil and Paraguay. Colours represent

714 different lineages.

715

716 **Figure S2. Fully annotated Brazilian SARS-CoV-2 time tree.** Time resolved maximum

717 likelihood phylogeny containing 17,135 high quality Brazilian SARS-CoV-2 near-full-genome

718 sequences (n=3,866 generated in this study) analysed against a backdrop of global reference

719 sequences. Variants of interest (VUM) and concern (VOC) are highlighted.

720

721 **Figure S3.** Number of total cases plotted against the number of sequences available for each  
722 location Brazilian state/region.

723

724 **Figure S4. Sensitivity of the viral introduction analysis to geographic sampling biases. A, B)**

725 Sensitivity analysis, by performing the discrete ancestral state reconstruction on 10 bootstrap

726 replicate trees, showing how the proportion of international imports and exports into and from

727 Brazil from external locations did not vary in the time and source of viral transitions. **C-K)**

728 Equivalent sensitivity analysis showing how the number of viral exchanges within Brazilian

729 regions by counting the state changes from the root to the tips did not vary in the time and source

730 of viral transitions.

731

732 **Figure S5. Sensitivity of the Gamma viral introduction analysis to geographic sampling**

733 **biases. A, B)** Sensitivity analysis, by performing the discrete ancestral state reconstruction on 10

734 bootstrap replicate trees, showing how the proportion of international imports and exports of

735 Gamma variant into and from Brazil from external locations did not vary in the time and source

736 of viral transitions. **C-K)** Equivalent sensitivity analysis showing how the number of exchanges

737 of the Gamma variant between Brazilian regions did not vary in the time and source of viral

738 transitions.

739

740 **Figure S6. Sensitivity of the Zeta viral introduction analysis to geographic sampling biases.**

741 **A, B)** Sensitivity analysis, by performing the discrete ancestral state reconstruction on 10

742 bootstrap replicate trees, showing how the proportion of international imports and exports of

743 Zeta variant into and from Brazil from external locations did not vary in the time and source of



744 viral transitions. **C-K)** Equivalent sensitivity analysis showing how the number of exchanges of  
745 the Zeta variant between Brazilian regions did not vary in the time and source of viral transitions.

746

747 **Figure S7. Temporal Signal.** (A) Root-to-tip regression plots for Brazilian lineages of concern  
748 and of interest (P.1 – Gamma and P.2- Zeta).

749

750 **Supplementary Table 1.** Information on the n=3803 SARS-CoV-2 Brazilian samples sequenced  
751 as part of this study.

752

753 **Supplementary Table 2.** Information on the n=63 SARS-CoV-2 from Paraguay samples  
754 sequenced as part of this study.

755 **Supplementary Table 3.** GISAID acknowledgment table.

756

## 757 **FIGURE LEGENDS**

758 **Fig. 1. Key events following the first confirmed infection of SARS-CoV-2 in Brazil.** (A)

759 Timeline of SARS-CoV-2 key events in Brazil. The Brazilian map was colored according to  
760 geographical macro region: North (red), Northeast (green), Southeast (purple), Midwest (light

761 blue) South (light orange). (B) Epidemic curve showing the progression of reported daily viral  
762 infection numbers in Brazil from the beginning of the epidemic (grey) and deaths (red) in the

763 same period, with restriction phases indicated along the bottom. (C) Map of cumulative SARS-  
764 CoV-2 cases per 100,000 inhabitants in Brazil up to June 2021.

765

766 **Fig. 2. Phylogenetic analysis and SARS-CoV-2 lineage dynamics in Brazil.** (A) Map of Brazil  
767 with the number of sequences in GISAID as of 30th June 2021. (B) Temporal sampling of  
768 sequences in Brazilian states through time with VOCs highlighted and annotated according to  
769 their PANGO lineage assignment. (C) Time resolved maximum likelihood phylogeny containing  
770 high quality near-full-genome sequences from Brazil (n=3866) obtained from this study, analysed  
771 against a backdrop of global reference sequences (n=25,288). Variants under monitoring (VUM)  
772 and concern (VOC) are highlighted on the phylogeny. (D) Sources of viral introductions into  
773 Brazil characterized as external introductions from the rest of the world. (E) Sources of viral  
774 exchanges (imports and exports) in and outside Brazil. (F) Number of viral exchanges within  
775 Brazilian regions by counting the state changes from the root to the tips of the phylogeny in  
776 panel C.

777

778 **Fig. 3. Spatiotemporal spread of VOC and VUM in Brazil.** (A) Phylogeographic  
779 reconstruction of the spread of the Gamma VOC in Brazil. Circles represent nodes of the  
780 maximum clade credibility phylogeny and are colored according to their inferred time of  
781 occurrence. Shaded areas represent the 80% highest posterior density interval and depict the  
782 uncertainty of the phylogeographic estimates for each node. Differences in population density  
783 are shown on a dark-white scale; B) Phylogeographic reconstruction of the spread of the Zeta  
784 VUM across Brazil. Circles represent nodes of the maximum clade credibility phylogeny and are  
785 colored according to their inferred time of occurrence. Shaded areas represent the 80% highest  
786 posterior density interval and depict the uncertainty of the phylogeographic estimates for each  
787 node. Differences in population density are shown on a dark-white scale; In both Panels (A) and  
788 (B) solid curved lines denote the links between nodes and the directionality of movement is anti-

789 clockwise along the curve (as shown in the “dispersal direction” sublegends). (C) Number of  
790 exchanges of the Gamma variant between Brazilian regions (N=North; NE=Northeast;  
791 MD=Midwest; SE=Southeast; S=South); (D) Number of exchanges of the Zeta variant between  
792 Brazilian regions; (E) Sources of viral export of the VOC and VUM from Brazil to the rest of the  
793 world.

794

795 **Fig. 4. The SARS-CoV-2 epidemic in Paraguay.** (A) Epidemic curve showing the progression  
796 of reported viral infection numbers in Paraguay from the beginning of the epidemic (grey) and  
797 deaths (red) in the same period; (B) Progressive distribution of the top 20 PANGO lineages in  
798 Paraguay over time; (C) Time resolved maximum likelihood tree containing (n=63) high quality  
799 near-complete genome sequences from Paraguay obtained in this study analysed against a  
800 backdrop of global reference sequences. VUM and VOCs are highlighted on the phylogeny.  
801 Small circles indicate genome sequences from Brazil. Bigger circles indicate sequences from  
802 Paraguay. New genomic sequences from Paraguay obtained in this study are highlighted with a  
803 red border.

26 Feb  
1st confirmed case  
in South America,  
61 yo male returning  
traveller from Italy

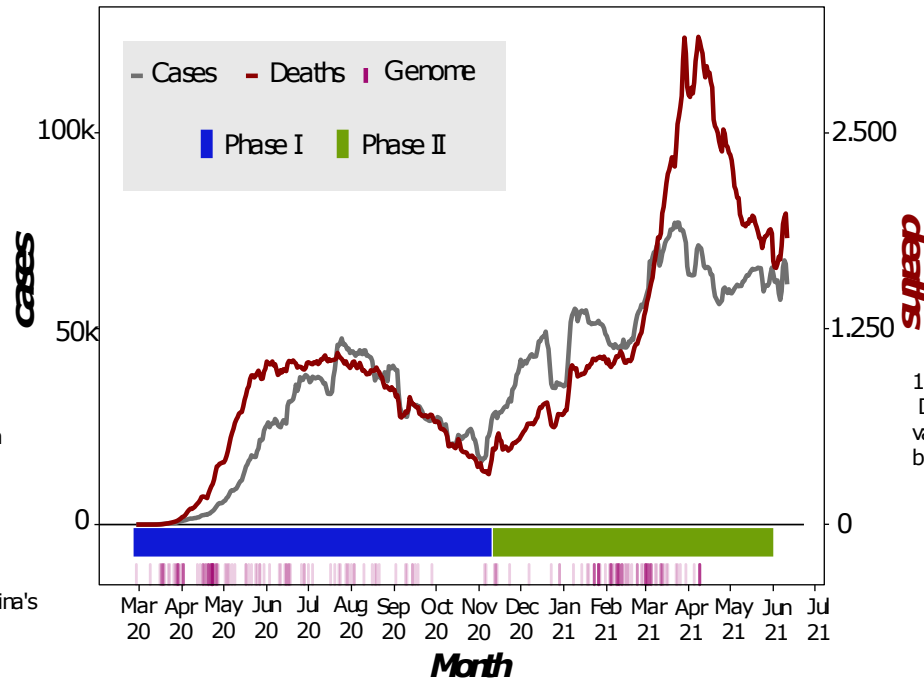
A

10 April  
The virus had reached  
remote locations;  
a Yanomami teen died  
of it in Roraima

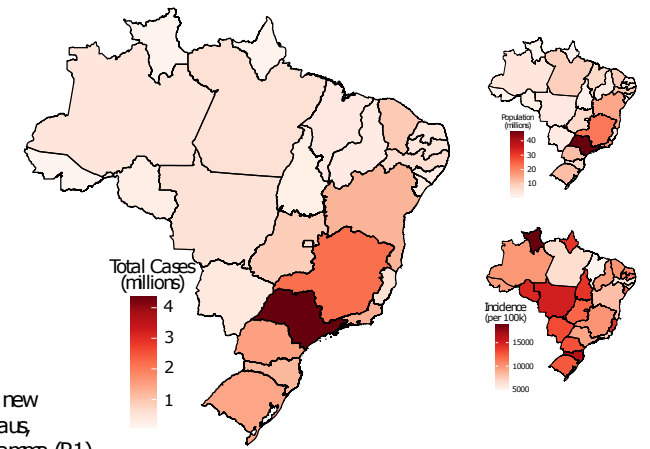
20 April  
Several cities started  
to ease social isolation  
guidelines in favor  
of contact trading.

30 April  
Brazil overtook China's  
official number of  
confirmed cases,  
surpassing 87,000

B



C



13 January:  
Discovery of a new  
variant in Manaus,  
being called Gamma (P.1)

17 January:  
Anvisa has authorized the  
emergency use of CoronaVac  
and the AstraZeneca vaccines

14 May  
Delta (B.1.617.2) Variant is identified  
in São Paulo



2020

2021

21 Mar  
All Brazilian states reported  
at least one confirmed  
case of COVID-19, BR-MoH  
declares large-scale community  
transmission in the country

17 Mar  
1st confirmed death for  
COVID-19, 61 yo male

11 Mar  
SARS-CoV-2 confirmed as pandemic  
by World Health Organization

Oct-Nov  
Zeta (P.2) Variant is identified  
in Rio de Janeiro

31 December:  
Alpha (B.1.1.7) Variant is identified  
in São Paulo

7 April:  
For the first time,  
Brazil has recorded  
over 4,000 confirmed  
COVID-19 deaths in 24 hours

04 April  
Beta (B.1.351) Variant is identified  
in São Paulo

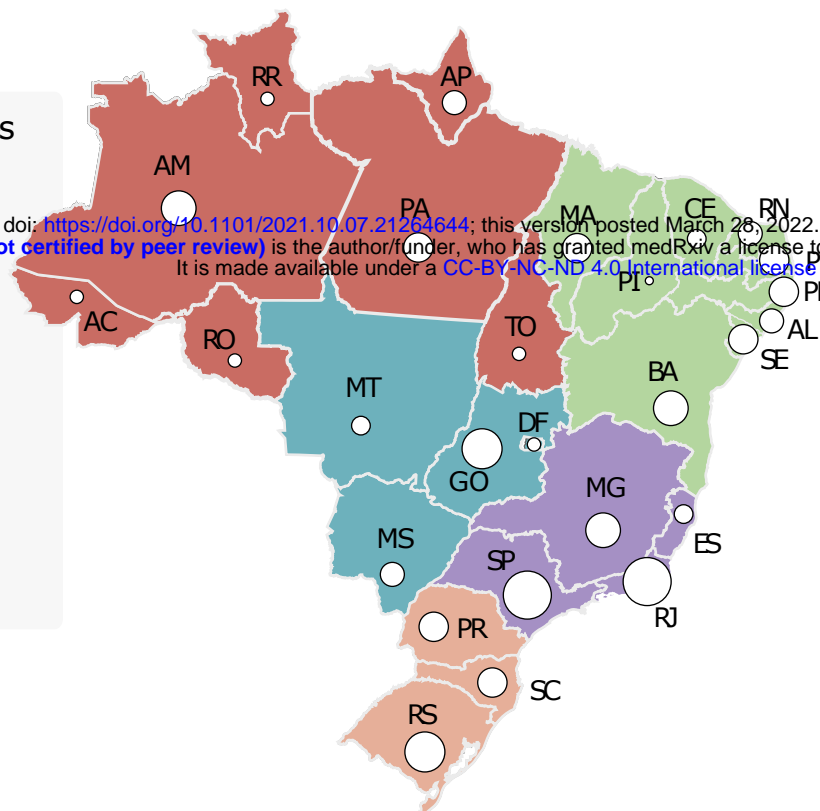
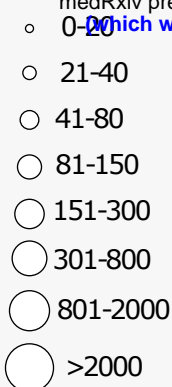
9 June:  
Brazil registered  
50 M people vaccinated

**Legend**

- Restriction measures
- Vaccination campaign
- First confirmed case
- First Brazilian VOI detected
- First confirmed death
- First Brazilian VOC detected

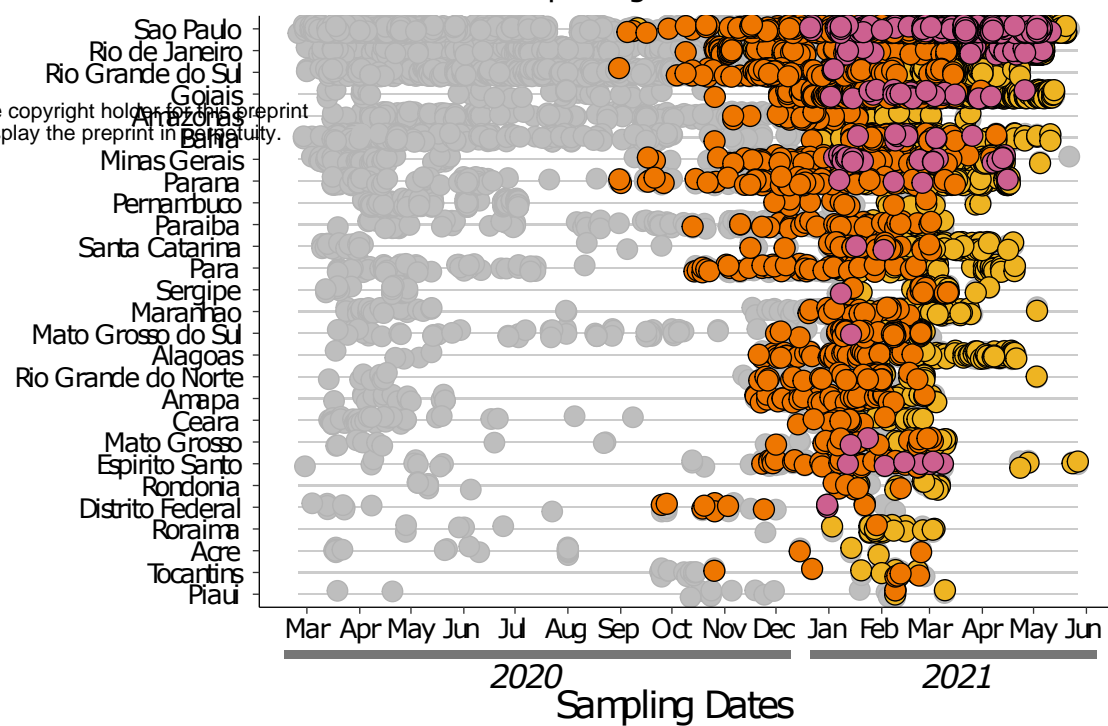
A

SARS-CoV-2 genomes from the Brazilian epidemic

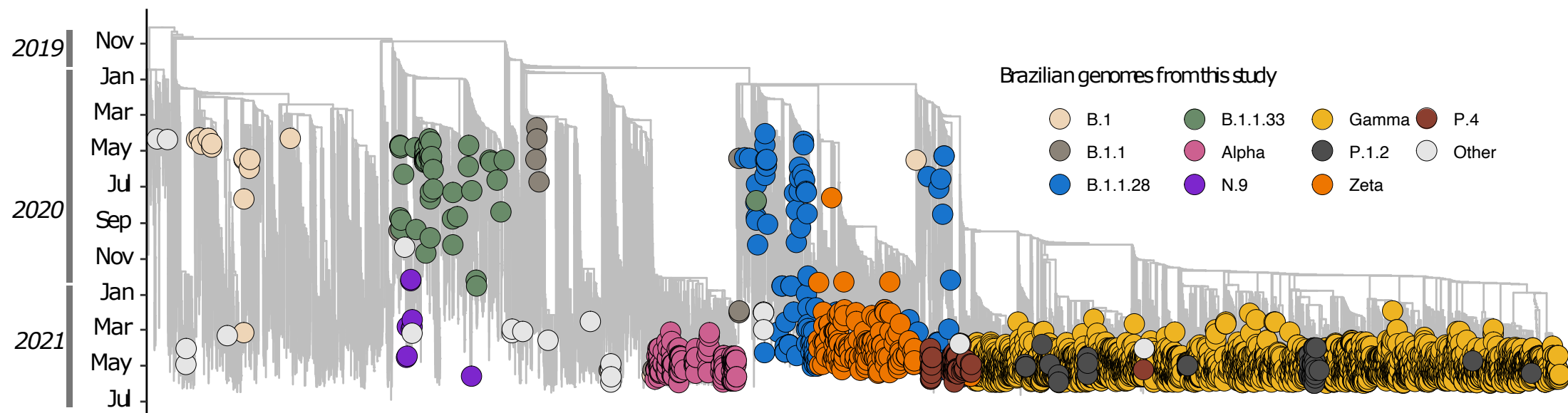


B

VOCs - VOIs Alpha Gamma Zeta Other Lineages  
Brazilian states with sequencing data

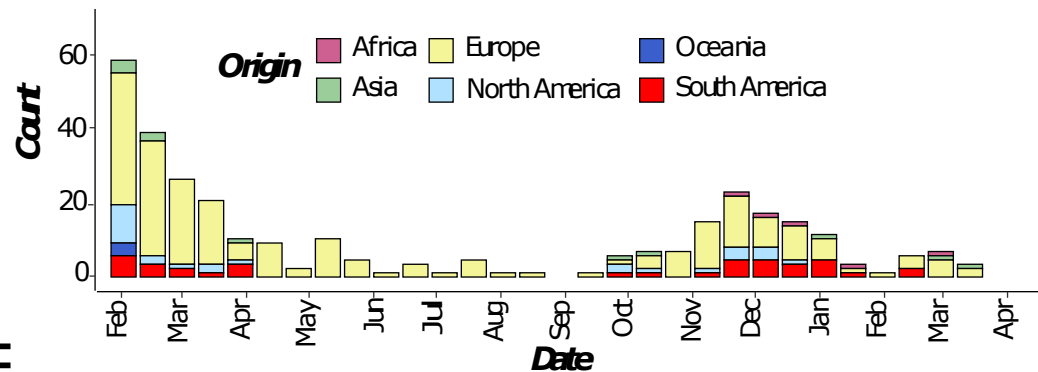


C



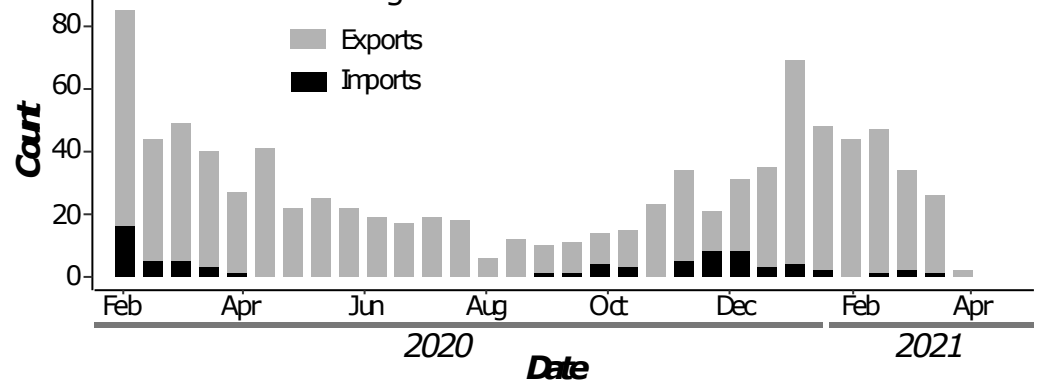
D

International introductions into Brazil



E

Source of viral exchanges



F

