Article

Genomic Epidemiology of SARS-CoV-2 in Tocantins State and the Diffusion of P.1.7 and AY.99.2 Lineages in Brazil

Ueric José Borges de Souza [1,*], Raíssa Nunes dos Santos [1], Fernando Lucas de Melo [2], Aline Belmok [2], Jucimária Dantas Galvão [3], Tereza Cristina Vieira de Rezende [3], Franciano Dias Pereira Cardoso [3], Rogério Fernandes Carvalho [4], Monike da Silva Oliveira [4,5], Jose Carlos Ribeiro Junior [4], Evgeni Evgeniev Gabev [6], Ester Cerdeira Sabino [7], Clarice Weis Arns [8], Bergmann Morais Ribeiro [2], Fernando Rosado Spilki [9] and Fabrício Souza Campos [1,*]

[1] Bioinformatics and Biotechnology Laboratory, Campus of Gurupi, Federal University of Tocantins, Gurupi 77410-570, Brazil; engraissanunes@gmail.com

[2] Baculovirus Laboratory, Department of Cell Biology, Institute of Biological Sciences, University of Brasilia, Brasília 70910-900, Brazil; flucasmelo@gmail.com (F.L.d.M.); abelmokadias@gmail.com (A.B.); bergmann.ribeiro@gmail.com (B.M.R.)

[3] Central Public Health Laboratory of the State of Tocantins, Palmas 77054-970, Brazil; jucydg@mail.uft.edu.br (J.D.G.); tereza.rezende@gmail.com (T.C.V.d.R.); francianocardoso@yahoo.com.br (F.D.P.C.)

[4] Laboratory of Microbiology, Federal University of North Tocantins, Araguaína 77804-970, Brazil; fcrogeriofc@hotmail.com (R.F.C.); monikeoliveira@discente.ufg.br (M.d.S.O.); ribeirojuniorjc@gmail.com (J.C.R.J.)

[5] Molecular Biology Laboratory, Institute of Biological Sciences, Federal University of Goiás, Goiânia 74001-970, Brazil

[6] Department of Physiology and Pathophysiology, Medical University of Sofia, 1431 Sofia, Bulgaria; egabev@medfac.mu-sofia.bg

[7] Department of Infectious and Parasitic Diseases, Institute of Tropical Medicine, Faculty of Medicine, University of São Paulo, São Paulo 05403-000, Brazil; sabinoec@gmail.com

[8] Laboratory of Animal Virology, Institute of Biology, University of Campinas, Campinas 13083-862, Brazil; clarns@gmail.com

[9] Molecular Microbiology Laboratory, Feevale University, Novo Hamburgo 93525-075, Brazil; fernandors@feevale.br

* Correspondence: uericjose@gmail.com (U.J.B.d.S.); camposvet@gmail.com (F.S.C.)

Abstract: Tocantins is a state in the cross-section between the Central-West, North and Northeast regions of Brazilian territory; it is a gathering point for travelers and transportation from the whole country. In this study, 9493 genome sequences, including 241 local SARS-CoV-2 samples (collected from 21 December 2020, to 16 December 2021, and sequenced in the MinION platform) were analyzed with the following aims: (i) identify the relative prevalence of SARS-CoV-2 lineages in the state of Tocantins; (ii) analyze them phylogenetically against global SARS-CoV-2 sequences; and (iii) hypothesize the viral dispersal routes of the two most abundant lineages found in our study using phylogenetic and phylogeographic approaches. The performed analysis demonstrated that the majority of the strains sequenced during the period belong to the Gamma P.1.7 (32.4%) lineage, followed by Delta AY.99.2 (27.8%), with the first detection of VOC Omicron. As expected, there was mainly a dispersion of P.1.7 from the state of São Paulo to Tocantins, with evidence of secondary spreads from Tocantins to Goiás, Mato Grosso, Amapá, and Pará. Rio de Janeiro was found to be the source of AY.99.2 and from then, multiple cluster transmission was observed across Brazilian states, especially São Paulo, Paraiba, Federal District, and Tocantins. These data show the importance of trade routes as pathways for the transportation of the virus from Southeast to Northern Brazil.

Keywords: SARS-CoV-2; variants of concern; genome analysis; viral evolution; viral trajectory

1. Introduction

The ongoing COVID-19 pandemic is one of the greatest global threats in modern history. Since 2019, with the host spillover of the Severe Acute Respiratory Syndrome Coronavirus 2

(SARS-CoV-2) [1], the novel betacoronavirus has had a severe impact on human health, the economy, and social life. This global scenario demonstrated the limitations of One Health policies in terms of minimizing transmission across continents. The lack of measures to control the new virus in Brazil included inadequate public interventions, non-scientific communications, social inequality, and slow vaccination [2,3]. Moreover, global transit networks increase the risk of importing and exporting pathogens, as was evidenced by the circulation of variants of concern (VOCs) around the world during the pandemic [4,5].

International public efforts have been focused on the emergence of virus variants leading to changes in viral fitness. The mutations present in SARS-CoV-2 variants may increase transmissibility, enhance escape from the human immune response, or alter biologically important phenotypes in a way that confers a fitness advantage, such as the mutations of the spike affecting antigenicity [6]. The global emergence of several VOCs, recently renamed by the World Health Organization (WHO), including the UK variant (Alpha or B.1.1.7), the South Africa variant (Beta or B.1.351), the Brazil variant (Gamma or P.1), and the India variant (Delta or B.1.617.2), highlights the utmost importance of genomic surveillance for monitoring viral mutations [7,8]. Gamma and Delta evolved into new sublineages, called P1.X and AY.X, respectively [9,10].

Brazil faced one of the highest numbers of confirmed SARS-CoV-2 cases worldwide, 27,968,811 with an incidence of 14,005.5 cases per 100,000 inhabitants (Brazilian Ministry of Health, https://covid.saude.gov.br, accessed on 15 March 2022). Tocantins is a state located in a very peculiar region of Brazil; it is a gathering point for transportation routes from the Southeast, Central-Western, North, and Northeast regions of the country (Figure 1). Tocantins presents two of the most impacted Brazilian biomes, namely Cerrado (88% of the state area) and Amazon (12%), and has been highlighted as the most recent deforestation hotspot related to soy and livestock production [11]. Soybean cultivation is spread throughout Tocantins, which favors the movement of people. In addition, the proximity of Tocantins state to the Amazonas state facilitated the spread of the Gamma variant [3]. The first patient infected with SARS-CoV-2 in Tocantins was reported on 16 March 2020. Since then, the state has reported 300,957 cases and 4132 deaths, with an incidence of 19,134.30 cases per 100,000 inhabitants (Integra Saúde Tocantins, http://integra.saude.to.gov.br/covid19/InformacoesEpidemiologicas, accessed on 15 March 2022).

Genomic surveillance provides important clues as to virus–host dynamics [12,13]. Long-read sequencing was first used in arboviral genomic studies [14,15] and then quickly adapted for SARS-CoV-2 given the public health emergency. Thus, next-generation sequencing (NGS) technologies represent a powerful tool for tracing the origin, spread, and transmission chains of outbreaks, as well as monitoring the evolution of etiological agents. The COVID-19 pandemic has triggered efforts for real-time surveillance strategies based on viral genome sequencing. In Brazil, the Corona-ômica-BR project (a network made up of 11 research institutions and dozens of researchers) is harnessing efforts and human resources to track the viral spread and support public health authorities. In this study, we report the complete genome sequences and phylogenetic analysis of 241 SARS-CoV-2 genomes detected in the Tocantins state, located in the North region of Brazil, from December 2020 to December 2021. Furthermore, the first description of SARS-CoV-2 molecular epidemiology in this Brazilian state is presented.
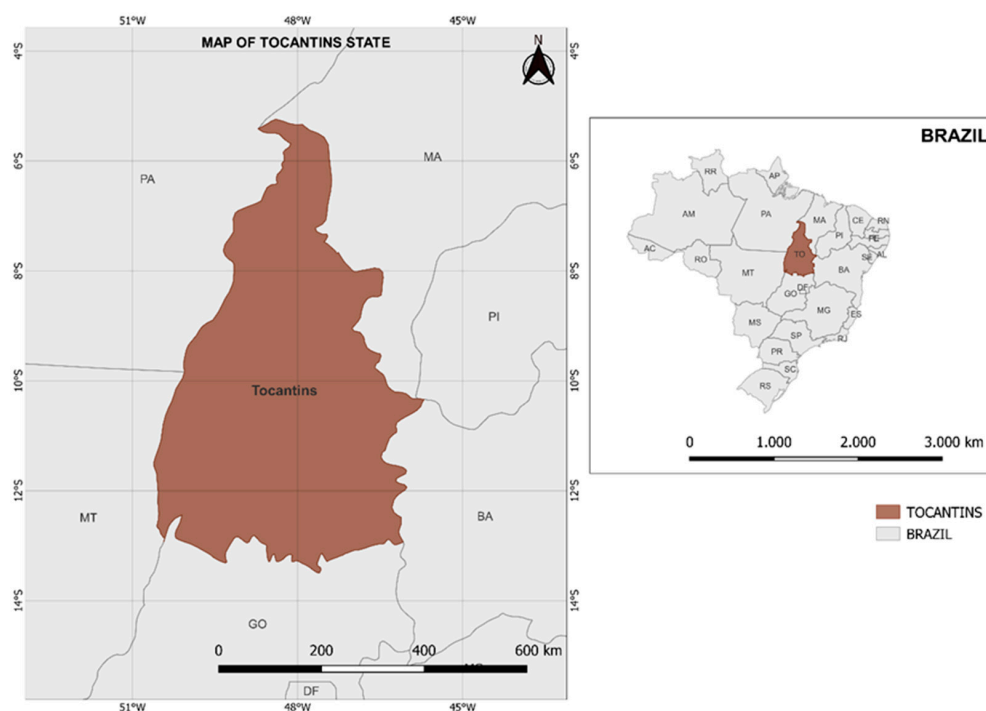
Figure 1. State map of Brazil with emphasis on Tocantins state. This figure was generated using QGIS v 3.20.2-Odense (Quantum Geographic Information System, QGIS Application Network Connections, Saint-Pierre, La Réunion, France).

2. Materials and Methods

2.1. Sample Collection

Samples from patients with SARS-CoV-2-positive nasopharyngeal RT-qPCR were collected between 21 December 2020, and 16 December 2021, at the Central Public Health Laboratory of the State of Tocantins (Figure 1). Reverse-transcription-quantitative real-time polymerase chain reaction (RT-qPCR) for SARS-CoV-2 is the gold-standard method for laboratory diagnosis of COVID-19. The kit used for all samples was the Allplex 2019-nCoV assay, Seegene Inc., Seoul, Korea (provided by the Brazilian Ministry of Health and used as a routine procedure at all public health laboratories in Brazil). The RT-qPCR reactions were conducted using 6.5 µL of RNA input, instead of the recommended 8 µL. Briefly, the Allplex kit employs target genes and labeled probes E (FAM), N (Quasar 670), RdRP (Cal Red 610), and an internal control (IC; HEX) in one multiplex RT-qPCR. All RT-qPCR reactions were performed using the QuantStudio 5 Real-Time PCR System (Applied Biosystems, Waltham, CA, USA). In total, 241 positive samples with quantification cycle (Cq) below 25 for at least one primer were selected and submitted to the SARS-CoV-2 genome sequencing. The study was approved by the Ethics Committee (33202820.7.1001.5348).

2.2. RNA Extraction and Sequencing

The samples were processed at the Central Public Health Laboratory of the State of Tocantins, with an Extracta kit Viral RNA MVXA-P096 FAST (Loccus, Brazil) using an automated extractor (Extracta 96, Loccus, Brazil) following the manufacturer's guidelines. The cDNA synthesis was performed using Luna Script RT SuperMix (5×) (New England Biolabs, Ipswich, MA, USA). The synthesized cDNAs were used as templates for a multiplexed PCR reaction using the two non-overlapping primer pools provided by the ARTIC Network to generate ~400 bp amplicons tiled across the genome (V3 nCov-2019 primers) (ARTIC primer set). Amplicons from both primer pools were combined and purified with a 1× volume of Ampure XP beads (Beckman Coulter, Brea, CA, USA). The MinION library preparation was performed using the Ligation Sequencing kit SQK-LSK-109 and Natives Barcoding kits EXP-NBD104 and EXP-NBD114 (Oxford Nanopore, Oxford, UK). The result-

ing library was loaded on R9.4 Oxford MinION flowcells (FLO-MIN106) and sequenced using the MinION Mk1B device. ONT MinKNOW software was used to collect raw data. High-accuracy base-calling and quality control analyses were performed using Guppy (v6.0.1) and NanoPlot (v.1.33.0), respectively. Assembly of the high accuracy base called fastq files was performed using the nCoV-2019 novel coronavirus bioinformatics protocol (https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html, accessed on 12 January 2022) with Minimap2 [16] and Medaka (https://github.com/nanoporetech/medaka, accessed on 12 January 2022) for consensus sequence generation.

2.3. Comparative Genome Analysis

All sequences were analyzed with Nextclade v1.7.2 [17] (https://clades.nextstrain.org/, accessed on 12 January 2022) to determine the clade and the number of gap regions. Lineages were assigned to each genome using Pangolin v3.1.17 [18] (https://github.com/cov-lineages/pangolin, accessed on 12 January 2022). The sequences were grouped according to the two most common VOCs (Gamma and Delta) and aligned using MAFFT v7.490 [19], with the strain Wuhan-Hu-1 genome reference (accession number: NC_045512.2) as the first sequence in the alignment. Genome map and SNP histogram were generated for each grouped lineage using the msastats.py script, and plotAlignment and plotSNPHist functions [20] (https://github.com/laduplessis/SARS-CoV-2_Guangdong_genomic_epidemiology/, accessed on 12 January 2022).

In order to identify single-nucleotide polymorphisms (SNPs) and insertions/deletions (INDELs), the assembled genomes were mapped against the reference (NC_045512.2) separately according to grouped lineages by Snippy variant calling and core genome alignment v4.6.0 (https://github.com/tseemann/snippy, accessed on 12 January 2022). This pipeline uses FreeBayes v1.3.2 [21] to call variants and snpEff v5.0 [22] to annotate variants based on their genomic locations and predicts coding effects such as synonymous or non-synonymous amino acid replacement.

2.4. Phylogenetic Analysis

The 241 Tocantins sequences were combined with 9289 SARS-CoV-2 genome sequences to evaluate the phylogenetic relationships of the viruses circulating in a global context. Accordingly, all genome sequences and associated metadata deposited until 31 December 2021, were downloaded from GISAID (A total of 6,665,096). The first round of subsampling was performed by using Nextstrain's bioinformatics toolkit [23]. We selected 10,000 sequences collected from Brazil between 28 February 2020 to 31 December 2021, excluding data from Tocantins. Next, we selected 5000 sequences from South America collected between 28 February 2020 to 31 December 2021, excluding data we had already selected from Brazil. Finally, we selected 2000 sequences from Africa, Asia, China, Europe, North America, and Oceania, with the maximum date limited to 31 December 2021. For all the sequence selections, we limited the minimum-length parameter to 29,000 base pairs. These steps generated a random subsampling of 25,000 SARS-CoV-2 genomes. To reduce the dataset of genomes to allow feasible phylogenetic analysis, we then used the genome-sampler [24], which selects temporal, geographical, and diversity context samples from an available dataset containing focal sequences. Tocantins' sequences were defined as the focal sequence collection. These steps resulted in a dataset with 9075 sequences. We also downloaded from the GISAID database 211 sequences from Tocantins state and included three sequences from December 2019 collected in Wuhan, China (EPI_ISL_402125, EPI_ISL_406798, and EPI_ISL_434534). The resulting dataset has a total of 9530 sequences.

Multiple sequence alignment was performed using MAFFT with default settings and manually inspected using AliView v1.27 [25]. The maximum likelihood (ML) phylogenetic tree was built using IQ-TREE v2.1.2 [26]. The IQ-TREE2 analysis was performed under the generalized time-reversible (GTR) model of nucleotide substitution with empirical base frequencies (+F) plus FreeRate model (+R3), as selected by the ModelFinder software [27] and 1000 replicates of ultrafast bootstrapping (−B 1000) and SH-aLRT branch test

(−alrt 1000). The ML tree was inspected in TempEst v1.5.3 [28] to investigate the temporal signal through a regression analysis of root-to-tip genetic distance against sampling dates. Thirty-seven sequences were strong outliers and were removed from the analysis, leaving 9493 sequences in the final dataset. By repeating the above steps, a strong temporal signal in the final dataset was confirmed (R2 of 0.83 and correlation coefficient of 0.91; Figure S1) and a dated phylogenetic tree was inferred using TreeTime [29]. Visualizations of the time-resolved phylogenetic tree were produced in R v4.1.2 using the ggtree package [30].

Additionally, we extracted sequences of two clades that grouped a large number of Tocantins sequences from the phylogenetic tree using the caper R package [31]. Multiple sequence alignment was performed using MAFFT with default settings and manually inspected using AliView. Spatial and temporal patterns of diffusion were estimated using a Bayesian Markov Chain Monte Carlo (MCMC) approach implemented in BEAST v1.10.4 [32] with BEAGLE v3.1.2 [33] to improve computational time. The HKY + G model [34] was used to model nucleotide evolution under a strict clock model with a Continuous Time Markov Chain (CTMC) [35] and a non-parametric Bayesian Sky-grid coalescent model [36]. The best-fit nucleotide substitution models for the two data sets were identified according to the Bayesian information criterion (BIC) method in jModelTest v2.1.10 [37]. Location diffusion rates were estimated using the Bayesian stochastic search variable selection (BSSVS) model [38] with a discretization scheme defined as Brazilian states. Following the Nextstrain workflow [23], the initial evolutionary clock rate was set to 8×10^{-4} substitutions per site per year. Two independent MCMC chains were performed with a length of 100 million generations sampling every 10,000 generations. The two independent runs were merged with Log Combiner v1.10.4 [32] and the convergence of the MCMC chain was assessed using Tracer v1.7.2 [39]. A maximum clade credibility tree (MCC) with annotated branches was then generated in TreeAnnotator v1.10.4, with 20% removed as burn-in [32]. The MCC phylogenetic trees were visualized using ggtree R package and the SpreaD3 v0.9.7.1 [40] was used to visualize the phylogeographic patterns embedded in MCC trees.

3. Results

Of the 241 samples from 21 December 2020 to 16 December 2021, 51.5% (*n* = 124) were from female patients, and the mean patient age was 40.3 years (interquartile range—IQR: 26 years). The mean cycle threshold (Ct) value for the RT-qPCR conducted at the Central Public Health Laboratory of the State of Tocantins was 19.3 cycles (min: 11.2, max: 26.8, median: 19.2; IQR: 3 cycles). In total, 241 samples from 46 cities from Tocantins state were used. The majority of the samples were from the capital, Palmas (*n* = 83; 34.4%), followed by Gurupi (*n* = 31; 12.9%), Araguaína (*n* = 18; 7.5%), Porto Nacional (*n* = 16; 6.6%), and Formoso do Araguaia (*n* = 8; 3.3%) (Table S1 and Figure S2).

3.1. Comparative Genome Analysis

In total, 13 Pango lineages were observed in the 241 sequenced genomes. The most frequently detected lineage was P.1.7 (78 sequences, 32.4%), followed by AY.99.2 (67 sequences, 27.8%), P.1 (56 sequences, 23.2%), AY.43 (27 sequences, 11.2%), and AY.42 (3 sequences, 1.2%). We also report here the first detection of the BA.1 lineage (VOC Omicron) in Tocantins, from a sample collected on 10 December 2021, in Gurupi municipality.

Considering the nomenclature proposed by the WHO, 136 genomes were classified as VOC Gamma and 102 genomes as VOC Delta. The genetic analysis revealed 502 different mutations comprising VOC Gamma samples, of which 52.8% (265) were non-synonymous (missense), 44.6% (224) were synonymous, 0.8% (4) were intergenic at 5′ untranslated region (UTR), 0.6% (3) were intergenic at 3′ untranslated region (UTR), and 1.2% (6) were in-frame deletions. The ORF1ab carried 62.4% (313) replacements followed by spike (S) (*n* = 73; 14.5%), ORF3a (*n* = 37; 7.4%), nucleocapsid (N) (*n* = 31; 6.2%), membrane (M) (*n* = 12; 2.4%), ORF8 (*n* = 11; 2.2%), ORF7a (*n* = 10; 2.0%), ORF7b (*n* = 6; 1.2%), ORF10 (*n* = 5; 1.0%), envelope (E) (*n* = 2; 0.4%), and ORF6 (*n* = 2; 0.4%). From the 502 different mutations,

30.7% (154) were identified in two or more sequences and 6.4% (32) were found in more than 50% of sequences (Figure 2). From these highly frequent mutations, 21 (65.6%) were non-synonymous, 9 (28.1%) were synonymous, 1 (3.1%) was intergenic at 5′ untranslated region (UTR), and 1 (3.1%) was an in-frame deletion. Among the non-synonymous mutations, 60.9% (13/21) were located in the S gene (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, P681H, T1027I, and V1176F).
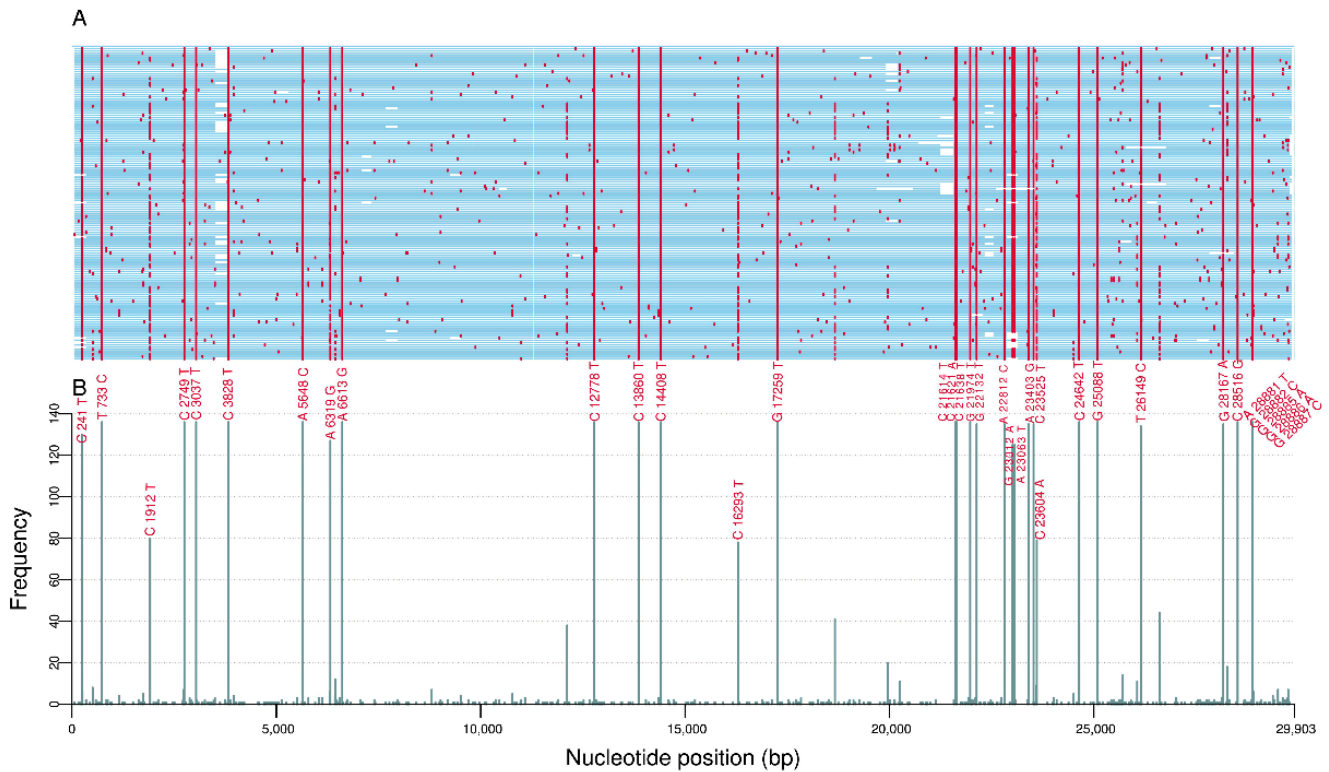

Figure 2. (**A**) Genomic location and (**B**) frequency of single-nucleotide polymorphisms (with respect to the reference genome NC_045512.2) among 136 sequences from VOC Gamma.

For the 102 sequences belonging the VOC Delta, 326 different mutations were found, of which 56.1% (183) were non-synonymous, 41.4% (135) were synonymous, 1.2% (4) were intergenic at 5′ untranslated region (UTR), 0.6% (2) were intergenic at 3′ untranslated region (UTR), and 0.6% (2) were in-frame deletions. Among these mutations, 60.7% (198) were found at ORF1ab followed by spike (*n* = 49; 15.0%), nucleocapsid (*n* = 26; 8.0%), ORF3a (*n* = 15; 4.6%), ORF7a (*n* = 14; 4.3%), ORF8 (*n* = 8; 2.5%), membrane (*n* = 7; 2.1%), envelope (*n* = 4; 1.2%), ORF7b (*n* = 3; 0.9%), and ORF10 (*n* = 2; 0.6%). From the total of mutations, 52.5% (171) were identified in two or more sequences. Highly frequent mutations (≥50% of genomes) were found in 34 genome positions (Figure 3). From these, 76.5% (26) mutations were non-synonymous, 11.8% (4) were synonymous, 5.9% (2) were intergenic at 5′ untranslated region (UTR), and 5.9% (2) were in-frame deletions. Remarkably, seven non-synonymous mutations were located in the S gene (T19R, G142D, L452R, T478K, D614G, P681R, and D950N).
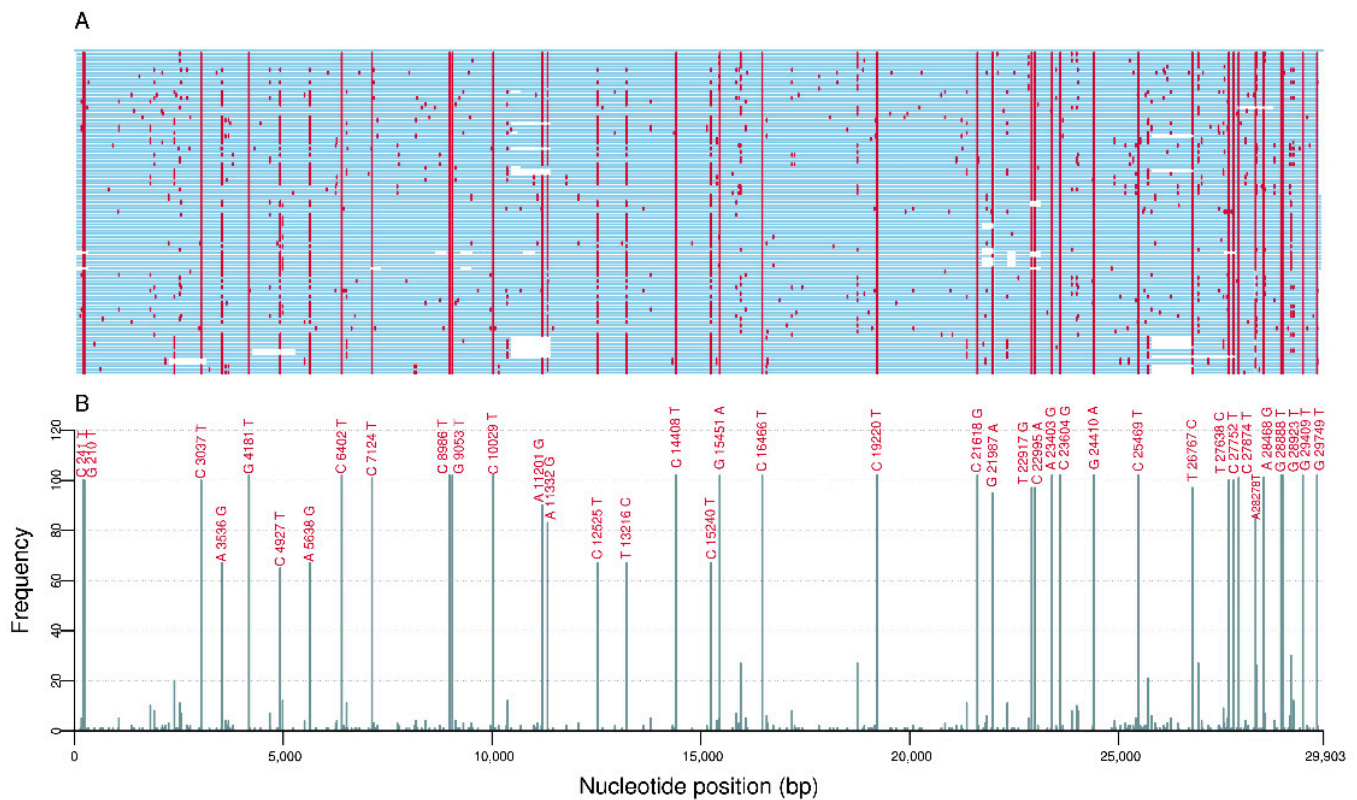
Figure 3. (**A**) Genomic location and (**B**) frequency of single-nucleotide polymorphisms (with respect to the reference genome NC_045512.2) among 102 sequences from VOC Delta.

Overall, the Tocantins comprised, until 31 December 2021, 452 SARS-CoV-2 genomes (241 sequenced in our study and 211 deposited in GISAID, until 31 December 2021). The genomic analysis with Pangolin revealed 21 SARS-CoV-2 lineages circulating in Tocantins state. The most abundant lineage was P.1, accounting for 33.4% (151 genomes) of the sequenced genomes. The second most common lineage was P.1.7 (27.0% or 122 genomes), followed by AY.99.2 (17.9% or 81 genomes), AY.43 (9.5% or 43 genomes), P.2 (3.5% or 16 genomes), and B.1.1.28 (3.1% or 14 genomes). The sequencing of SARS-CoV-2 genomes from Tocantins started in September 2020 (Figure 4). We point out here that since the emergence and spread of Gamma-related lineages (P.1 and its sub-lineages, especially P.1.7) in Brazil, they have quickly become the dominant SARS-CoV-2 lineages circulating in Tocantins state. This was a pattern clearly observed until August 2021, when the Delta lineage was introduced, and, thereafter, the sub-lineages AY.99.2, AY.43, and AY.42. During September 2021, 31 genomes were sequenced, of which 38.7% (12) belonged to the AY.99.2 lineage, followed by P.1.7 (29.0% or 9), AY.43 (9.7% or 3), and AY.42 (6.5% or 2). In October and November only, lineages associated with VOC Delta were detected among the sequenced genomes. Subsequently, in December, one genome of lineage BA.1 (VOC Omicron) was detected.
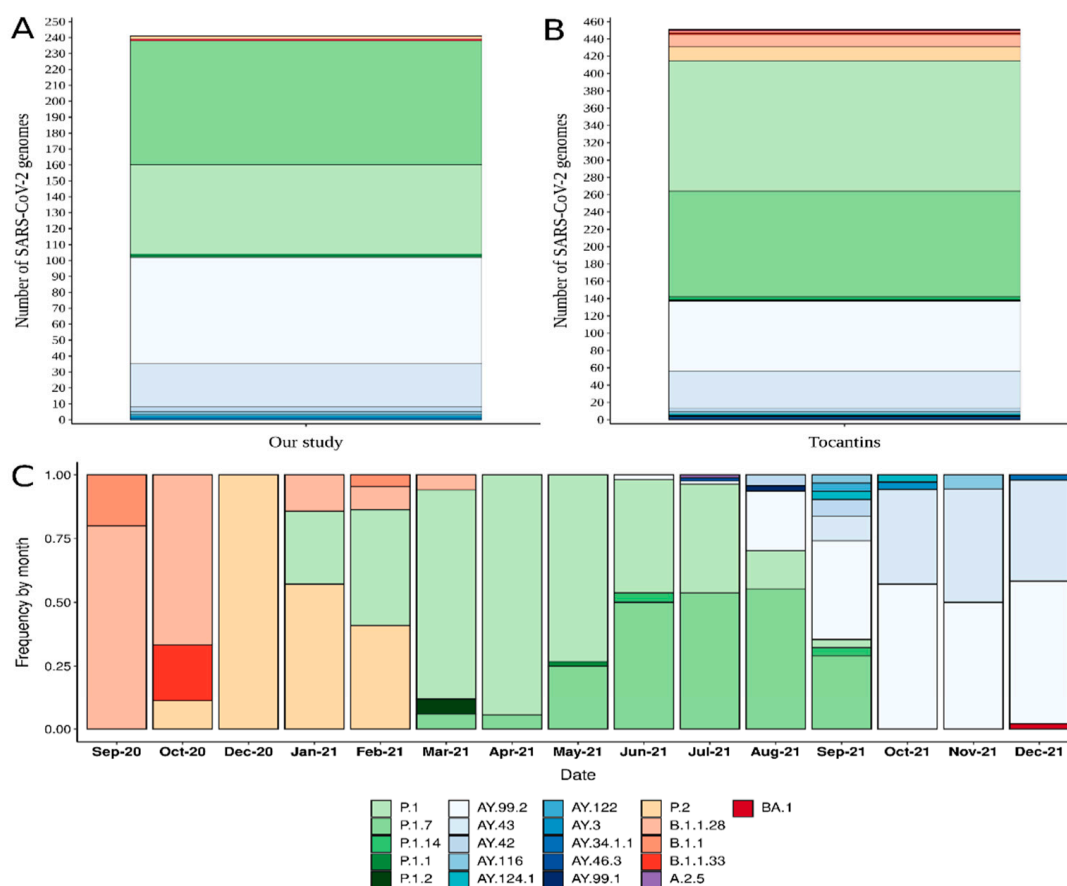
Figure 4. SARS-CoV-2 lineages circulating in Tocantins state. (**A**) Lineages identified in this study. (**B**) Overall lineages in the Tocantins. (**C**) Frequency of the most abundant lineages from September 2020 to December 2021.

3.2. Phylogenetic Analysis

The time-resolved phylogenetic tree confirmed the PANGO lineages assigned, and the sequences from the Tocantins were allocated into branches close to Brazilian and South American sequences (Figure 5). We also point out that during the initial period, the predominant lineages circulating in Brazil were B.1.1.28 and B.1.1.33, followed by the rise of P.2, a descendant of the B.1.1.28 lineage, which formed a clear, evolving differentiation. Furthermore, the first sequences obtained from the Tocantins were grouped in these clades as well, as expected. Subsequently, the majority of the analyzed samples were classified as belonging to the Brazilian P.1 lineage (Gamma), which has dominated the epidemiological scenario in Brazil since its emergence (November–December 2020). Additionally, P.1.7, which originated from the P.1 lineage (spike substitution P681H), was widely spread in Tocantins from March 2021 [8]. The first confirmed case of Delta lineage in Brazil occurred on 26 April 2021. In Tocantins state, the Delta VOC was first detected in samples collected on 29 June 2021, and became dominant in September 2021. We highlight here the high spread and circulation of the Delta variant, and, in particular, of the lineage AY.99.2 in Tocantins, as well as in all Brazilian regions (Figure 5; see also Figure S3). The only BA.1 sequence identified in this study branched in a clade represented by 59 sequences, mostly represented by sequences from Australia (37.3% or 22 genomes), followed by sequences from Brazil (13.6%, or eight genomes; six from São Paulo, one from Ceará, and one from the Tocantins) (Figure 5).
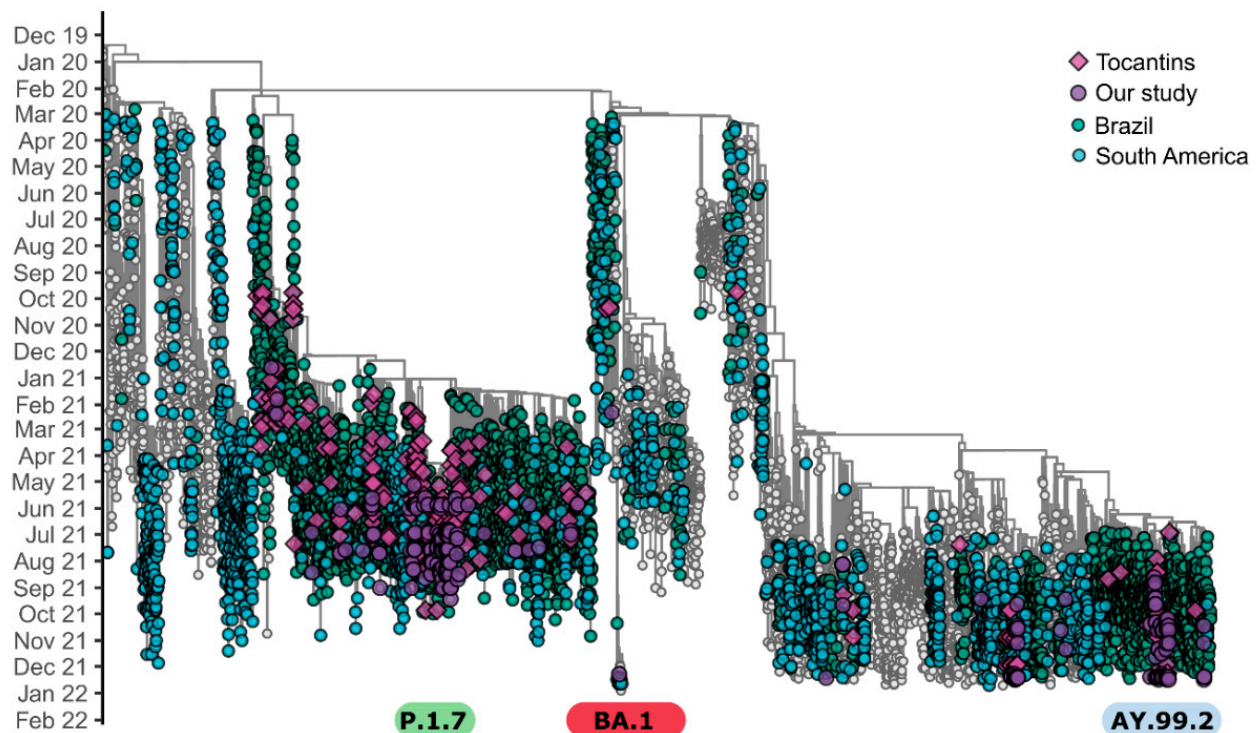
Figure 5. Time-resolved maximum-likelihood tree of 9493 SARS-CoV-2 genomes. Tips are colored by the most frequent Pango lineages. Diamonds represent Tocantins genomes.

We identified two subclades with high branch support, one for the P.1.7 lineage (SH-aLRT = 75.8) and one for the AY.99.2 lineage (SH-aLRT = 84.7), both grouping a large number of sequences from Tocantins, especially from our study. The P.1.7 clade was composed of 243 sequences, the majority of which were from Tocantins (49.4% or 120 genomes), followed by São Paulo (24.3% or 59 genomes) and Goiás (12.3% or 30 genomes). Furthermore, this clade was represented by 15 Brazilian states and one genome from Chile (EPI_ISL_2663617). The AY.99.2 clade was composed of 214 sequences. The majority of these were also from Tocantins (34.1% or 73 genomes). Moreover, this clade contained only Brazilian sequences, represented by 19 states and the Federal District. In order to determine the time of the most recent common ancestor (TMRCA) and the diffusion of these two subclades identified in our maximum-likelihood phylogenetic tree, we used coalescent and phylogeographic methods.

For the P.1.7 subclade, we estimated a median evolutionary rate of 4.91×10^{-4} (95% highest posterior density interval (HDP): 4.21×10^{-4} to 5.69×10^{-4} substitutions per site per year) and the TMRCA on 10 January 2021 (95% HPD: 27 October 2020 to 2 March 2021) (Figure 6A). Interestingly, the tree was rooted in a sample from Tocantins (EPI_ISL_2756458; PP: 1.0) and the BSSVS procedure identified well-supported rates of posterior diffusion from Tocantins to Goiás (BF: 32,149; PP: 0.99), Mato Grosso (BF: 17.54; PP: 0.55), Amapá (BF: 12.72; PP:0.47), and Pará (BF: 126.15; PP: 0.89). The tree reconstruction also showed important migrations from Tocantins to Rio Grande do Sul (BF: 18.33; PP: 0.51) and Pernambuco (BF: 10.51; PP: 0.42) (Figure 6B).
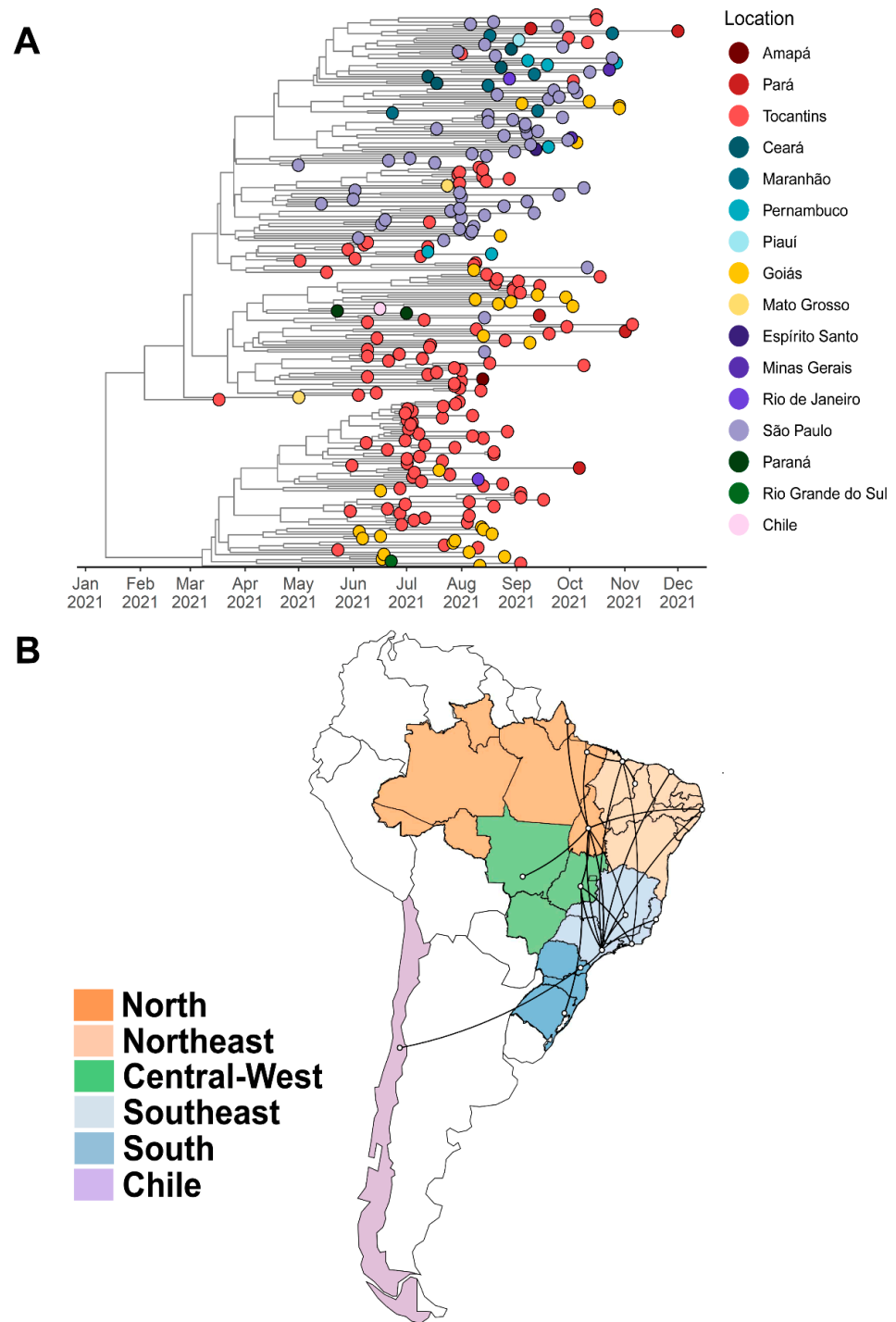
Figure 6. (**A**) Bayesian time-scale phylogeny of 243 P.1.7 samples circulating in Brazil from 13 March 2021, to 1 October 2021. Tips are colored based on sampling location, following the legend's colors on the right. The molecular clock was inferred using BEAST with the HKY + G evolution model and a strict molecular clock (initial evolutionary rate was set to 8×10^{-4} substitutions/site/year). (**B**) Spatiotemporal diffusion of the filtered P.1.7 subclade across the Brazilian states. The figure shows well-supported transition rates of the P.1.7 lineage between Brazilian regions in discrete phylogeographic reconstructions using BSSVS procedure and Bayes Factor tests. The following states belong to each Brazilian region: North: AC, AM, AP, PA, RO, RR, and TO. Northeast: AL, BA, CE, MA, PB, PE, PI, RN, and SE. Central-West: DF, GO, MS, and MT. South: PR, RS and SC. Southeast: ES, MG, RJ, and SP. The Tocantins state is located in the North region.

For the AY.99.2 clade, the TMRCA was dated 7 May 2021 (95% HPD: 26 March 2021 to 20 June 2021), and the median evolutionary rate was 6.20×10^{-4} (95% HPD: 5.37×10^{-4} to 7.20×10^{-4}). The state of Rio de Janeiro was identified the most probable source of its emergence (PP = 1.0), after which multiple cluster transmission was observed across the Brazilian states, especially São Paulo, Paraiba, Federal District, and Tocantins (Figure 7A). We also found evidence of large local transmission clusters of AY.99.2 lineage in Tocantins state and secondary disseminations from Tocantins to Pará (BF: 911.27; PP: 0.98). The tree reconstruction also showed important migrations from São Paulo (BF: 336.86; PP: 0.95) and Maranhão (BF: 46.33; PP: 0.71) to Tocantins (Figure 7B).
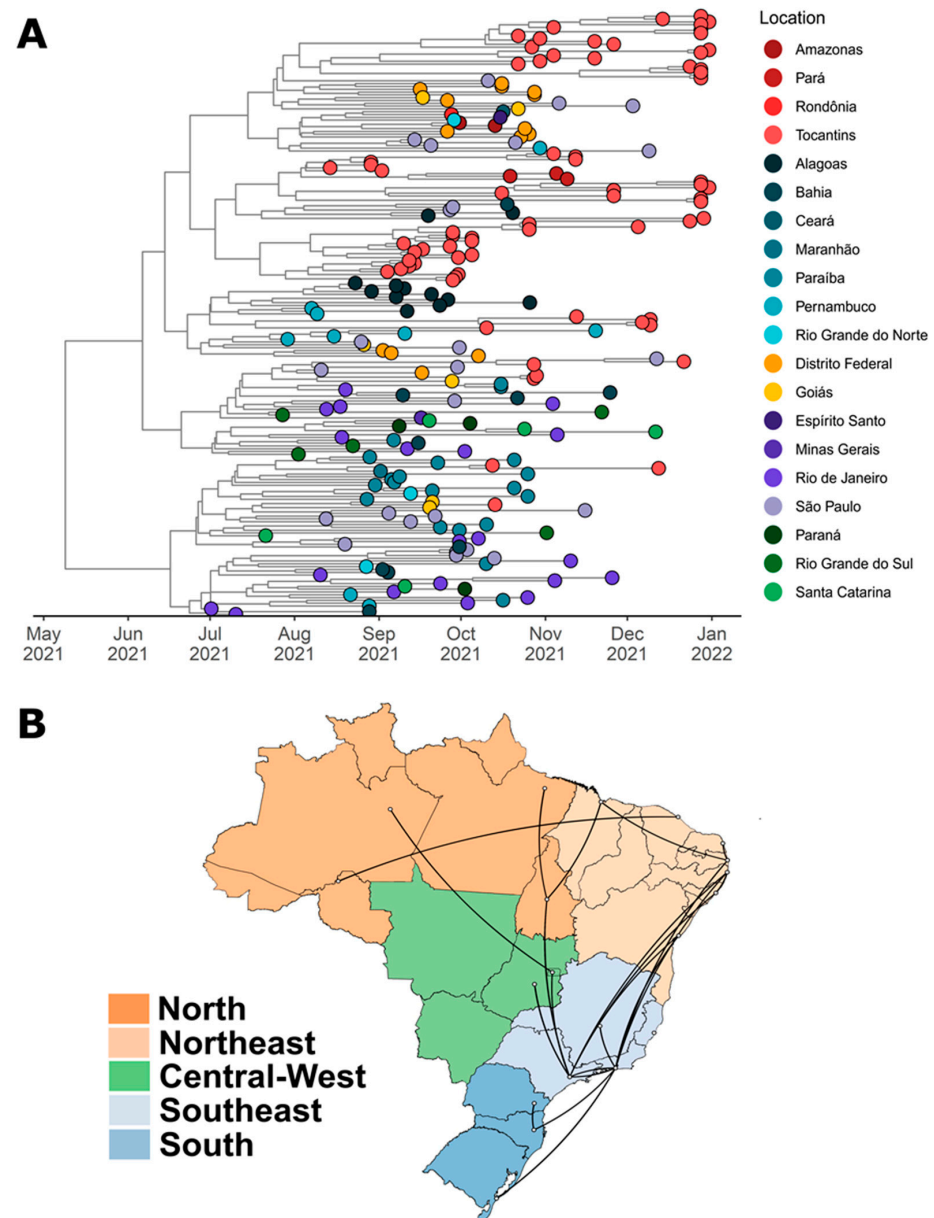
Figure 7. (**A**) Bayesian time-scale phylogeny of 214 AY.99.2 samples circulating in Brazil from 5 July 2021 to 16 December 2021. Tips are colored based on sampling location, following the legend's colors on the right. The molecular clock was inferred using BEAST with the HKY + G evolution model and a strict molecular clock (initial evolutionary rate was set to 8×10^{-4} substitutions/site/year). (**B**) Spatiotemporal diffusion of the filtered AY.99.2 subclade across the Brazilian states. The figure shows well-supported transition rates of the AY.99.2 lineage between Brazilian regions in discrete phylogeographic reconstructions using BSSVS procedure and Bayes Factor tests.

4. Discussion

In this study, we report the first description of SARS-CoV-2 molecular epidemiology in Tocantins State, Brazil, from September 2020 to December 2021. We demonstrate that three main lineages dominated the SARS-CoV-2 scenario in Tocantins during this period of time: P.1 (33.4%), P.1.7 (34.9%), and AY.99.2 (17.9%). Therefore, our analysis showed that in September 2020, SARS-CoV-2 strains were mainly classified as B.1.1.28 and B.1.1.33, which is similar to observations in other Brazilian regions [2,41,42]. Nevertheless, these lineages were rapidly replaced by the variant of interest (VOI) Zeta (P.2) and, especially, by the Gamma VOC (P.1 and P.1.7), which rapidly spread and dominated the epidemiological scenarios across different regions of Brazil [8]. The P.1 lineage was first detected on 6 January 2021, by Japan's National Institute of Infectious Disease, in four travelers who arrived in Tokyo from Amazonas, Brazil, on 2 January 2021, at airport control [3,43]. The rapid increase in the number of hospitalizations caused by this SARS-CoV-2 variant was a significant source of economic and social disruption in Brazil and led to a significant amount of pressure on the health system. AY.99.2 was a prominent ramification of VOC Delta in all of South America before the emergence of VOC Omicron [44].

The Delta VOC was first detected in the Indian state of Maharashtra in October 2020, with subsequent dissemination throughout India. The first confirmed case of Delta in Brazil occurred on 26 April 2021. A Delta lineage was first detected in Tocantins in one sample by the end of June 2021, representing 1.67% (1/60). In August 2021, it was identified in 29.8% (14/47), and in September, in more than 65% (15/23) of sequenced samples. During October and November, Delta variants became dominant in Tocantins. The observed growth in COVID-19 cases caused by Delta suggested a transition from individual to community transmission of this variant in August 2021 in Tocantins state.

The differences in viral kinetics found between the pre-Gamma, Gamma, and Delta variants in Tocantins suggest some incremental, but potentially adaptive, changes in viral dynamics. These changes are associated with the evolution of SARS-CoV-2 towards a more rapid viral spread. The emerging Brazilian Gamma variant, derived from the B.1.1.28 lineage, began spreading rapidly later in 2020 [45]. The Gamma variant spread from the north of Brazil, passing through Tocantins, towards the rest of the country. Moreover, the spread of Gamma was influenced by the demand for care for Amazon patients in other states that led to the transfer of patients to different hospitals throughout the country. Naveca et al. (2021) [46] reported that P.1+ lineages with convergent mutations evolved independently in other Brazilian states, and that the lineage designated as P.1.7 (P.1 + P681H) was the major lineage circulating outside Amazonas.

Our study indicates that the lineage AY.99.2 probably emerged in Rio de Janeiro state, and that multiple cluster transmission was observed across Brazilian states. We also point out a strong migration from Tocantins to Pará. All these observed migrations were probably influenced by the Belém–Brasília highway. In addition, the Tocantins sequences formed a separate clade, reinforcing the idea that local transmission occurred in the state. This may have been mainly influenced by the relaxation of non-pharmacological measures, such as social isolation and quarantines [46–48].

We recognize some limitations in the present study. First, the number of analyzed genomes corresponds to a small percentage of the number of cases. Second, the dispersal routes were influenced by multiple factors, including the geographic parameters and traffic metrics of people along the Belém–Brasília highway [49]. However, this may have been a determining factor in local transmission. Third, the dissemination of VOCs among Brazilian states is not a linear description [50,51]. The introduction of the Delta variant occurred in the first semester of 2021, during a period of restriction relaxation and incomplete vaccine schedules. Despite this, and despite the differences between states with large population concentrations (São Paulo, Rio de Janeiro), others, such as Tocantins, are influenced by the transport of agribusiness products [52,53]. In the case of the latter, the state of Tocantins stands out, as it is considered the last Brazilian agricultural frontier beyond the Amazon [11].

Our findings help to explain how the AY.99.2 lineage is transmitted so efficiently in Brazilian populations. Specifically, our results suggest that São Paulo, Federal District, Paraíba, and Tocantins states played an important role in this transmission in the Midwest, North, and Northeast regions of the country. This highlights that genomic surveillance, in addition to being useful in determining which viral variants are circulating in a given location, can also help elucidate viral trajectories. Our findings were also important to support actions to reinforce vaccination, epidemiological surveillance, and the planning of health units (intensive care units and inpatient beds), in addition to the structuring of the genomic surveillance department at the Central Public Health Laboratory of the State of Tocantins (LACEN-TO).

Despite the limitations of our study, we showed genomic surveillance to be a potential tool for monitoring the circulation of SARS-CoV-2 in Tocantins state. Although we used a relatively small number of SARS-CoV-2 sequences from Tocantins, we suggest that our analysis reflects the major variants circulating, as shown in other studies on other Brazilian states. However, levels of SARS-CoV-2 sequencing were significantly different between Brazil and other countries and across Brazilian regions and states [8]. For example, states from northern regions, such as Acre, Amapá, Roraima, and Tocantins, submitted 231, 517, 256, 452, and 890 sequences in GISAID until 31 December 2021, respectively. In comparison, states from regions in the southeast, such as São Paulo and Rio de Janeiro, submitted 45,434 and 11,081 sequences, respectively. This highlights the urgent need for implementing efficient strategies for genomic surveillance in Brazil across different regions, such as the action of the Corona-Ômica Network.Br-MCTI, allowing local public health agencies to know more about the pandemic situation in their states.

Finally, SARS-CoV-2 lineages presenting different kinetics are still emerging in the COVID-19 pandemic. We show the importance of the continuous monitoring of the epidemiology of COVID-19, as well as the specific studies of dynamics that are useful to defining public health measures. All viral surveillance information is deeply helpful in order to follow the evolution of the virus and its routes of dispersion.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Andersen, K.G.; Rambaut, A.; Lipkin, W.I.; Holmes, E.C.; Garry, R.F. The proximal origin of SARS-CoV-2. *Nat. Med.* **2020**, *26*, 450–452. [CrossRef] [PubMed]
2. Candido, D.S.; Claro, I.M.; de Jesus, J.G.; Souza, W.M.; Moreira, F.R.R.; Dellicour, S.; Mellan, T.A.; Du Plessis, L.; Pereira, R.H.M.; Sales, F.C.S.; et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science* **2020**, *369*, 55–60. [CrossRef] [PubMed]
3. Faria, N.R.; Mellan, T.A.; Whittaker, C.; Claro, I.M.; Candido, D.S.; Mishra, S.; Crispim, M.A.E.; Sales, F.C.S.; Hawryluk, I.; McCrone, J.T.; et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **2021**, *372*, 815–821. [CrossRef] [PubMed]
4. Baker, R.E.; Mahmud, A.S.; Miller, I.F.; Rajeev, M.; Rasambainarivo, F.; Rice, B.L.; Takahashi, S.; Tatem, A.J.; Wagner, C.E.; Wang, L.F.; et al. Infectious disease in an era of global change. *Nat. Rev. Microbiol.* **2021**, *13*, 193–205. [CrossRef]
5. Hodcroft, E.B.; Zuber, M.; Nadeau, S.; Vaughan, T.G.; Crawford, K.H.D.; Althaus, C.L.; Reichmuth, M.L.; Bowen, J.E.; Walls, A.C.; Corti, D.; et al. Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **2021**, *595*, 707–712. [CrossRef]
6. Long, S.W.; Olsen, R.J.; Christensen, P.A.; Subedi, S.; Olson, R.; Davis, J.J.; Saavedra, M.O.; Yerramilli, P.; Pruitt, L.; Reppond, K.; et al. Sequence Analysis of 20,453 Severe Acute Respiratory Syndrome Coronavirus 2 Genomes from the Houston Metropolitan Area Identifies the Emergence and Widespread Distribution of Multiple Isolates of All Major Variants of Concern. *Am. J. Pathol.* **2021**, *191*, 983–992. [CrossRef]
7. Umair, M.; Ikram, A.; Salman, M.; Khurshid, A.; Alam, M.; Badar, N.; Suleman, R.; Tahir, F.; Sharif, S.; Montgomery, J.; et al. Whole-genome sequencing of SARS-CoV-2 reveals the detection of G614 variant in Pakistan. *PLoS ONE* **2021**, *16*, e0248371. [CrossRef]
8. Souza, U.J.B.; dos Santos, R.N.; Campos, F.S.; Lourenço, K.L.; da Fonseca, F.G.; Spilki, F.R. High Rate of Mutational Events in SARS-CoV-2 Genomes across Brazilian Geographical Regions, February 2020 to June 2021. *Viruses* **2021**, *13*, 1806. [CrossRef]
9. CDC. Emerging SARS-CoV-2 Variants. Available online: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-emerging-variants.html (accessed on 2 February 2022).
10. WHO. Tracking SARS-CoV-2 Variants. Available online: https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/ (accessed on 2 February 2022).
11. Drost, S.; Kuepper, B.; Piotrowski, M. Brazilian State of Tocantins: A Hotspot for Cerrado Deforestation. Chain Reaction Research. 2019, pp. 1–15. Available online: https://chainreactionresearch.com/wp-content/uploads/2019/04/Tocantins-Report-2.pdf (accessed on 13 November 2021).
12. Esper, F.P.; Cheng, Y.-W.; Adhikari, T.M.; Tu, Z.J.; Li, D.; Li, E.A.; Farkas, D.H.; Procop, G.W.; Ko, J.S.; Chan, T.A.; et al. Genomic Epidemiology of SARS-CoV-2 Infection During the Initial Pandemic Wave and Association with Disease Severity. *JAMA Netw. Open* **2021**, *4*, e217746. [CrossRef]
13. Meredith, L.W.; Hamilton, W.L.; Warne, B.; Houldcroft, C.J.; Hosmillo, M.; Jahun, A.S.; Curran, M.D.; Parmar, S.; Caller, L.G.; Caddy, S.L.; et al. Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: A prospective genomic surveillance study. *Lancet Infect. Dis.* **2020**, *20*, 1263–1271. [CrossRef]
14. Quick, J.; Grubaugh, N.D.; Pullan, S.T.; Claro, I.M.; Smith, A.D.; Gangavarapu, K.; Oliveira, G.; Robles-Sikisaka, R.; Rogers, T.F.; A Beutler, N.; et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* **2017**, *12*, 1261–1276. [CrossRef]
15. Vairo, F.; Coussoud-Mavoungou, M.P.; Ntoumi, F.; Castilletti, C.; Kitembo, L.; Haider, N.; Carletti, F.; Colavita, F.; Gruber, C.; Iannetta, M.; et al. Chikungunya Outbreak Group Taskforce. Chikungunya Outbreak in the Republic of the Congo, 2019 —Epidemiological, Virological and Entomological Findings of a South-North Multidisciplinary Taskforce Investigation. *Viruses* **2020**, *12*, 1020. [CrossRef] [PubMed]
16. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [CrossRef] [PubMed]
17. Aksamentov, I.; Roemer, C.; Hodcroft, E.B.; Neher, R.A. Nextclade: Clade assignment, mutation calling and quality control for viral genomes. *J. Open Source Softw.* **2021**, *6*, 3773. [CrossRef]
18. Rambaut, A.; Holmes, E.C.; O'Toole, Á.; Hill, V.; McCrone, J.T.; Ruis, C.; du Plessis, L.; Pybus, O.G. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **2020**, *5*, 1403–1407. [CrossRef]
19. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]
20. Lu, J.; du Plessis, L.; Liu, Z.; Hill, V.; Kang, M.; Lin, H.; Sun, J.; François, S.; Kraemer, M.U.; Faria, N.R.; et al. Genomic Epidemiology of SARS-CoV-2 in Guangdong Province, China. *Cell* **2020**, *181*, 997–1003.e9. [CrossRef]
21. Garrison, E.; Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv* **2012**, arXiv:1207.3907. [CrossRef]

22. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **2012**, *6*, 80–92. [CrossRef]

23. Hadfield, J.; Megill, C.; Bell, S.M.; Huddleston, J.; Potter, B.; Callender, C.; Sagulenko, P.; Bedford, T.; Neher, R.A. Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* **2018**, *34*, 4121–4123. [CrossRef]

24. Bolyen, E.; Dillon, M.R.; Bokulich, N.A.; Ladner, J.T.; Larsen, B.B.; Hepp, C.M.; Lemmer, D.; Sahl, J.W.; Sanchez, A.; Holdgraf, C.; et al. Reproducibly sampling SARS-CoV-2 genomes across time, geography, and viral diversity. *F1000Research* **2020**, *9*, 657. [CrossRef] [PubMed]

25. Larsson, A. AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **2014**, *30*, 3276–3278. [CrossRef] [PubMed]

26. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [CrossRef] [PubMed]

27. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermiin, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Meth.* **2017**, *14*, 587–589. [CrossRef] [PubMed]

28. Rambaut, A.; Lam, T.T.; Max Carvalho, L.; Pybus, O.G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2016**, *2*, vew007. [CrossRef]

29. Sagulenko, P.; Puller, V.; Neher, R.A. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol.* **2018**, *4*, vex042. [CrossRef]

30. Yu, G.; Smith, D.K.; Zhu, H.; Guan, Y.; Lam, T.T. ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **2017**, *8*, 28–36. [CrossRef]

31. Orme, D.; Freckleton, R.; Thomas, G.; Petzoldt, T.; Fritz, S.; Isaac, N.; Pearse, W. Caper: Comparative analyses of phylogenetics and evolution in R. R package version 05. *Proc. R. Soc. B* **2012**, *2*, 458.

32. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2018**, *4*, vey016. [CrossRef]

33. Ayres, D.L.; Darling, A.; Zwickl, D.J.; Beerli, P.; Holder, M.T.; Lewis, P.O.; Huelsenbeck, J.P.; Ronquist, F.; Swofford, D.L.; Cummings, M.P. BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics. *Syst. Biol.* **2012**, *61*, 170–173. [CrossRef]

34. Hasegawa, M.; Kishino, H.; Yano, T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **1985**, *22*, 160–174. [CrossRef] [PubMed]

35. Ferreira, M.A.R.; Suchard, M.A. Bayesian analysis of elapsed times in continuous-time Markov chains. *Can. J. Stat.* **2008**, *36*, 355–368. [CrossRef]

36. Gill, M.S.; Lemey, P.; Faria, N.R.; Rambaut, A.; Shapiro, B.; Suchard, M.A. Improving Bayesian Population Dynamics Inference: A Coalescent-Based Model for Multiple Loci. *Mol. Biol. Evol.* **2013**, *30*, 713–724. [CrossRef] [PubMed]

37. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef] [PubMed]

38. Lemey, P.; Rambaut, A.; Drummond, A.J.; Suchard, M.A. Bayesian Phylogeography Finds Its Roots. *PLoS Comput. Biol.* **2009**, *5*, e1000520. [CrossRef] [PubMed]

39. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [CrossRef] [PubMed]

40. Bielejec, F.; Baele, G.; Vrancken, B.; Suchard, M.A.; Rambaut, A.; Lemey, P. SpreaD3: Interactive Visualization of Spatiotemporal History and Trait Evolutionary Processes. *Mol. Biol. Evol.* **2016**, *33*, 2167–2169. [CrossRef]

41. Franceschi, V.B.; Caldana, G.D.; de Menezes Mayer, A.; Cybis, G.B.; Neves, C.A.M.; Ferrareze, P.A.G.; Demoliner, M.; de Almeida, P.R.; Gularte, J.S.; Hansen, A.W.; et al. Genomic epidemiology of SARS-CoV-2 in Esteio, Rio Grande do Sul, Brazil. *BMC Genom.* **2021**, *22*, 371. [CrossRef]

42. Slavov, S.N.; Giovanetti, M.; dos Santos Bezerra, R.; Fonseca, V.; Santos, E.V.; Rodrigues, E.S.; Adelino, T.; Xavier, J.; Borges, J.S.; Evaristo, M.; et al. Molecular surveillance of the on-going SARS-COV-2 epidemic in Ribeirao Preto City, Brazil. *Infect. Genet. Evol.* **2021**, *93*, 104976. [CrossRef]

43. Fujino, T.; Nomoto, H.; Kutsuna, S.; Ujiie, M.; Suzuki, T.; Sato, R.; Fujimoto, T.; Kuroda, M.; Wakita, T.; Ohmagari, N. Novel SARS-CoV-2 Variant in Travelers from Brazil to Japan. *Emerg. Infect. Dis.* **2021**, *27*, 1243. [CrossRef]

44. Gularte, J.S.; da Silva, M.S.; Mosena, A.C.S.; Demoliner, M.; Hansen, A.W.; Filippi, M.; Pereira, V.M.d.A.G.; Heldt, F.H.; Weber, M.N.; de Almeida, P.R.; et al. Early introduction.; dispersal and evolution of Delta SARS-CoV-2 in Southern Brazil, late predominance of AY.99.2 and AY.101 related lineages. *Virus Res.* **2022**, *311*, 198702. [CrossRef] [PubMed]

45. Nonaka, C.K.V.; Gräf, T.; Barcia, C.A.L.; Costa, V.F.; de Oliveira, J.L.; Passos, R.D.H.; Bastos, I.N.; de Santana, M.C.B.; Santos, I.M.; de Sousa, K.A.F.; et al. SARS-CoV-2 variant of concern P.1 (Gamma) infection in young and middle-aged patients admitted to the intensive care units of a single hospital in Salvador, Northeast Brazil, February 2021. *Int. J. Infect. Dis.* **2021**, *111*, 47–54. [CrossRef] [PubMed]

46. Naveca, F.G.; Nascimento, V.; de Souza, V.C.; Corado, A.L.; Nascimento, F.; Silva, G.; da Costa, C.; Duarte, D.; Pessoa, K.; Mejía, M.; et al. COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence. *Nat. Med.* **2021**, *27*, 1230–1238. [CrossRef] [PubMed]

47. Sabino, E.C.; Buss, L.F.; Carvalho, M.P.S.; Prete, C.A., Jr.; Crispim, M.A.E.; Fraiji, N.A.; Pereira, R.H.M.; Parag, K.V.; da Silva Peixoto, P.; Kraemer, M.U.G.; et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* **2021**, *397*, 452–455. [CrossRef]

48. Sonabend, R.; Whittles, L.K.; Imai, N.; Perez-Guzman, P.N.; Knock, E.S.; Rawson, T.; Gaythorpe, K.A.M.; Djaafara, B.; Hinsley, W.; FitzJohn, R.G.; et al. Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: A mathematical modelling study. *Lancet* **2021**, *398*, 1825–1835. [CrossRef]

49. Moran, E.F. Roads and Dams: Infrastructure-Driven Transformations in the Brazilian Amazon. *Ambiente Soc.* **2016**, *19*, 207–220. [CrossRef]

50. Naveca, F.; Nascimento, V.; Souza, V.; de Lima Corado, A.; Nascimento, F.; Mejía, M.; Brandão, M.J.; da Silva, A.V.; Benzaken, A.S.; Silva, G.; et al. The SARS-CoV-2 Variant Delta Displaced the Variants Gamma and Gamma Plus in Amazonas, Brazil. Available online: https://virological.org/t/the-sars-cov-2-variant-delta-displaced-the-variants-gamma-and-gamma-plus-in-amazonas-brazil/765 (accessed on 14 November 2021).

51. Patané, J.; Viala, V.; Lima, L.; Martins, A.; Barros, C.; Marqueze, E.; Bernardino, J.; Moretti, D.; Slavov, S.; Bezerra, R.; et al. SARS-CoV-2 Delta variant of concern in Brazil—Multiple introductions, communitary transmission, and early signs of local evolution. *medRxiv* **2021**. [CrossRef]

52. Albuquerque, M.G.; Santos, J.R.; Silveira, A.F.; Bornia Junior, D.L.; Nunes, R.B.; Gandra, T.B.R.; da Silva, A.L.C.; de Miranda, G.B.; Trevisol, A.R. Influence of socio-economic indicators and territorial networks at the spatiotemporal spread dynamics of COVID-19 in Brazil. *Soc. Nat.* **2021**, *33*, e59688. [CrossRef]

53. Castro, R.R.; Santos, R.S.C.; Sousa, G.J.B.; Pinheiro, Y.T.; Martins, R.R.I.M.; Pereira, M.L.D.; Silva, R.A.R. Spatial dynamics of the COVID-19 pandemic in Brazil. *Epidemiol. Infect.* **2021**, *149*, e60. [CrossRef]