

RESEARCH ARTICLE

Genomic characterization of Sabiá virus in Brazil, 2019–2020: Implications for diagnostics, virus evolution, and receptor binding

Ingra M. Claro^{1,2,3,4}, Erika R. Manuli^{1,2,5}, Camila A. M. da Silva^{1,2}, Thaís M. Coletti^{1,2}, Philippe Lemey⁶, Ana Catharina Nastro⁷, Luciana Vilas Boas Casadio⁷, Amaro Nunes Duarte-Neto^{8,9}, Joshua Quick¹⁰, Camila M. Romano^{1,2}, Charles Whittaker³, Sarah C. Hill^{11,12}, Carlos A. Prete¹³, Darlan S. Candido³, Filipe R. R. Moreira^{3,14}, Mariana S. Ramundo^{1,2,15,16}, Ian Nunes Valença^{1,2,3}, Jaqueline G. de Jesus^{1,2}, Flavia C. S. Sales^{1,2}, Mariana S. Cunha^{1,2,17}, Juliana M. Guerra⁹, Maria Cassia Mendes-Correa^{1,2}, Tania R. Tozetto-Mendoza², Marcilio Jorge Fumagalli¹⁸, Yeh-Li Ho⁷, Peter Simmonds¹², Weng M. Ng^{19,20,21}, Thomas A. Bowden²¹, William M. de Souza⁴, Oliver G. Pybus^{11,12}, Anna S. Levin^{1,2,22}, Nicholas Loman^{10†*}, Ester C. Sabino^{1,2,5†*}, Nuno R. Faria^{1,2,3†*}



1 Departamento de Moléstias Infecciosas e Parasitárias, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, **2** Instituto de Medicina Tropical, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, **3** School of Public Health, Imperial College London, MRC Centre for Global Infectious Disease Analysis and the Abdul Latif Jameel Institute for Disease and Emergency Analytics (J-IDEA), London, United Kingdom, **4** Department of Microbiology, Immunology and Molecular Genetics, College of Medicine, University of Kentucky, Lexington, Kentucky, United States of America, **5** Universidade Municipal de São Caetano do Sul, São Caetano do Sul, São Paulo, Brazil, **6** Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven – University of Leuven, Leuven, Belgium, **7** Divisão de Doenças Infecciosas, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil, **8** Departamento de Patologia, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, **9** Núcleo de Anatomia Patológica, Instituto Adolfo Lutz, São Paulo, Brazil, **10** School of Biosciences, University of Birmingham, Birmingham, United Kingdom, **11** Department of Biology, University of Oxford, Oxford, United Kingdom, **12** Department of Pathobiology and Population Sciences, Royal Veterinary College, Hatfield, United Kingdom, **13** Departamento de Engenharia de Sistemas Eletrônicos, Escola Politécnica da Universidade de São Paulo, São Paulo, Brazil, **14** Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, **15** Laboratório de Imunologia, INCOR, Hospital das Clínicas HCFMUSP, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil, **16** Departamento de Clínica Médica, Disciplina de Imunologia Clínica e Alergia, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, **17** Centro de Virologia, Núcleo de Doenças de Transmissão Vetorial, Instituto Adolfo Lutz, São Paulo, Brazil, **18** Centro de Pesquisa em Virologia, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, Brazil, **19** Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot, United Kingdom, **20** Electron Bio-Imaging Centre, Diamond Light Source, Didcot, United Kingdom, **21** Division of Structural Biology, Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, **22** Department of Infection Control, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil

OPEN ACCESS

Citation: Claro IM, Manuli ER, da Silva CAM, Coletti TM, Lemey P, Nastro AC, et al. (2026) Genomic characterization of Sabiá virus in Brazil, 2019–2020: Implications for diagnostics, virus evolution, and receptor binding. *PLoS Negl Trop Dis* 20(2): e0014008. <https://doi.org/10.1371/journal.pntd.0014008>

Editor: Eric Mossel, Center for Disease Control and Prevention, UNITED STATES OF AMERICA

Received: May 14, 2025

Accepted: February 6, 2026

Published: February 20, 2026

Copyright: © 2026 Claro et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The complete genome sequences of sample SP17 have been deposited in GenBank under accession numbers PX308626 and PX308627. The partial genome sequences of sample SP19, together with the arenavirus alignments used in this

☞ These authors contributed equally to this work.

‡ These senior authors contributed equally to this work.

* nfaria@ic.ac.uk (NRF); sabinoec@usp.br (ECS); nick@loman.net (NL)

Abstract

Between December 2019 and January 2020, two patients suspected of having severe yellow fever were admitted to a tertiary healthcare facility in São Paulo, Brazil, presenting with acute hemorrhagic syndrome and neurological alterations; both cases had fatal outcomes. Upon admission, both tested negative for yellow fever

study, are publicly available at https://github.com/CADDE-CENTRE/SABV_Brazil.

Funding: This work was supported by a Medical Research Council-São Paulo Research Foundation (FAPESP) CADDE partnership award (MR/S0195/1 and FAPESP 2018/14389-0 to ECS, NRF, NJL, and OGP), and by the São Paulo Research Foundation (FAPESP; grants 2017/08012-9 and 2023/11521-3 to IMC, 2018/17176-8 and 2019/12000-1 to JGJ, 2018/25468-9 to FCSS, 2019/07544-2 to TMC, 2019/21301-5 and 2021/10241-1 to CAMS, 2022/10408-6 to CMR). Additional funding was provided by the Bill & Melinda Gates Foundation (INV-034540 to IMC, ECS and NRF), the Wellcome Trust (DeZi project to 316633/Z/24/Z to NRF, ARTIC network 206298 to NJL and JQ; Early Career Award 226997/Z/23/Z to WMN; Digital Technology Development Award 226075/Z/22/Z to WMdS; 220414/Z/20/Z to SCH; and Wellcome Centre grant 203141/Z/16/Z to TAB), the UK Medical Research Council (MR/S007555/1 to TAB), and the European Research Council (725422-ReservoirDOCS to PL). The study also received funding from the European Union's Horizon 2020 Programme (MOOD 874850 to PL), the Research Foundation – Flanders (G0D5117N, G0B9317N, and G051322N to PL), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; grants to ASL and ECS), the Oxford Martin School (to OGP), the Global Virus Network Fellowship (to WMdS), and the Faculdade de Medicina da Universidade de São Paulo (to MSR). This study was funded by the Medical Research Council (MRC) Centre for Global Infectious Disease Analysis (MR/X020258/1) funded by the UK MRC and carried out in the frame of the Global Health EDCTP3 Joint Undertaking supported by the EU. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

viral RNA, and Sabiá virus (SABV), a New World arenavirus, was identified as the causative pathogen. To date, only four humans naturally acquired SABV infections have been confirmed, all fatal and linked to rural settings. We applied next-generation sequencing to generate complete and near-complete genomes from two patients (SP17 and SP19). Existing molecular diagnostics failed to detect SABV; therefore, new molecular tests were developed. Genetic analyses of SP17 and SP19 genomes along with other arenaviruses, revealed that the new cases were genetically diverse, showing 93-98.2% amino acid identity at the NP level among SP17, SP19, and the 1990 reference strain (SPH114202). Time-scaled phylogenetic analyses confirmed that SP17 and SP19 were not epidemiologically linked and suggested that SABV has been circulating undetected in Brazil for over a century. Additionally, homology modeling and structure-based mapping provided insights into SABV receptor-binding sequence conservation, suggesting that SABV shares similar receptor binding structure to other clade B arenaviruses, despite some amino acid variation around receptor binding site. Our findings underscore the need for retrospective and prospective surveillance of undiagnosed hemorrhagic fever cases to assess the public health impact of SABV in Brazil.

Author summary

Sabiá virus (SABV), also known as *Brazilian mammarenavirus*, causes severe hemorrhagic fever, with mortality rates among New World arenaviruses averaging around 30%. Over the past 40 years, four naturally acquired cases have been confirmed in Brazil, all fatal and linked to rural settings. Here, we characterized the genomic features of the third and fourth naturally acquired cases, identified in 2019 and 2020 in São Paulo, Brazil, in patients initially suspected of yellow fever. These cases were detected using untargeted metagenomic sequencing in a São Paulo healthcare facility. Our findings reveal that existing molecular diagnostics failed to detect contemporary SABV strains, leading to the development of novel assays based on newly generated SABV genomes. Genetic analyses confirm that these cases were not epidemiologically linked and provide evidence of SABV circulation in Brazil for over a century. Additionally, we present new insights into SABV receptor interactions, improving our understanding of its tropism. These findings expand the limited knowledge of SABV epidemiology and genetic diversity, with direct implications for diagnostics, surveillance, and outbreak preparedness against this highly pathogenic virus. Greater efforts should focus on pathogen-agnostic investigations of hemorrhagic fever cases, particularly in patients who test negative for yellow fever. Expanded surveillance is essential to better understand the disease spectrum, pathogenicity, transmissibility, and zoonotic reservoir of SABV.

Introduction

Sabiá virus (SABV) is an enveloped negative-stranded, bi-segmented RNA virus that is classified as a *Brazilian mammarenavirus*, *Arenaviridae* family [1]. Within the family, SABV belongs to Clade B from the New World arenaviruses complex (NWA) [2]. This clade contains other highly pathogenic threats known to cause severe hemorrhagic fever disease, including Chaparé (CHHF), Machupo (MACV), Junín (JUNV), and Guanarito (GTOV) viruses [2]. Like other mammarenaviruses, SABV contains two single-stranded, ambisense RNA segments: the small (S) segment (~3,500 nt) encodes the envelope glycoprotein precursor (GPC) and the nucleoprotein (NP), while the large (L) segment (~7,200 nt) encodes the zinc-binding matrix protein (Z) and the viral RNA-dependent RNA polymerase (RdRp) [3].

SABV was first identified in 1990 in the municipality of Cotia, São Paulo state, Brazil [2]. Nearly a decade later, in 1999, a second human infection was identified in Espírito Santo do Pinhal, about 155 km north of Cotia, also in São Paulo state [4]. Additionally, two laboratory-associated infections occurred following accidental aerosol exposure, one in 1992 (Brazil) and another in 1994 (USA). The 1992 case recovered with supportive care alone, without antiviral treatment, whereas the 1994 case was treated with ribavirin and recovered; however, a causal role of ribavirin in clinical recovery cannot be established based on this single case [5,6].

Our understanding of SABV evolution, pathogenicity, and reservoir remains limited due to insufficient data, partly resulting from the lack of specific diagnostic tools [7] and associated biohazard risks. Work with this virus requires biosafety level 4 (BSL-4) facilities, which are currently unavailable in South America. Although the tropism of viruses within the NWA complex is varied, MACV, JUNV, GTOV and SABV seem to utilize the human transferrin receptor (hTfR1) for entry into host cells [8]. Arenaviruses are primarily transmitted through aerosols, with wild rodents acting as the principal reservoirs. Human infection typically occurs via inhalation of aerosols contaminated with rodent excreta [9].

In this study, we investigated two SABV infections identified in patients hospitalized with severe hemorrhagic fever initially suspected of yellow fever at São Paulo's main tertiary hospital between December 2019 and January 2020. Using viral genomic data generated directly from clinical samples, we examine the evolutionary history of contemporary SABV strains and assess whether genetic divergence from the historical reference strain has implications for molecular detection and receptor usage.

Results

Untargeted metagenomic sequencing using the SMART-9N metagenomic approach, confirmed the presence of SABV RNA in patient SP17 [10,11], a 52-year-old male from Sorocaba, SP, Brazil, with a history of forest hiking who sought care at a basic health facility in SP on December 30, 2019. His condition worsened, leading to referral to the *Hospital das Clínicas* on January 6, 2020, where the patient died on January 11, 2020.

Sequencing of seven additional YFV-negative retrospective cases enabled the identification of a previously undetected SABV case, SP19, a 63-year-old male rural-worker from Assis, São Paulo, Brazil, who was admitted to the *Hospital das Clínicas* on December 10, 2019. SP19 reported no travel within two weeks before symptom onset. SP19 experienced a 9-day illness (Table A in S1 Text). SP19 died on December 12, 2019, with a diagnosis of sepsis in the death certificate (Fig 1 and Table A in S1 Text). Both SABV's positive cases were reported to São Paulo Secretary of Health. Details of infection location, symptom onset, and death dates for SP17, SP19, and previously detected Case 1 (1990) and Case 2 (1999) are shown in Fig 1 and Table B in S1 Text. Additional details of clinical manifestations and histopathology were reported elsewhere [10]. Additionally, Patient SP4 was diagnosed with hepatitis A infection, with 99.65% of genome identity with a strain from São Paulo (Accession number: MG049743.1) (Fig A and Tables A and C in S1 Text). The other five patients were discharged and a viral cause for disease was not identified (Table A in S1 Text). These findings highlight that multiple viral pathogens may be responsible for hemorrhagic syndromes in YFV-negative cases.

To ensure the reliability of the SABV findings and eliminate the possibility of between-sample or barcode contamination, the two putative-SABV samples were re-sequenced using new flow cells (one per replicate). Samples from SP17 and SP19 yielded 14,423,445 and 6,424,328 reads, respectively, of which 0.12% and 0.002% aligned to the SABV genome

virus strain names and countries of identification for clade B NWA viral species. Scale bar=nucleotide substitutions per site (s/s), H=hemorrhagic, L=large, S=small, bp=base pairs, BR=Brazil, AR=Argentina, BO=Bolivia, VE=Venezuela. Maps in panels A and B were created in R version 4.0.1 using the *ggplot2*, *sf*, *maturalearth*, *geobr*, *tmap*, *lwgeom*, *units*, and *mapview* packages. Basemap data: Natural Earth (<http://www.naturalearthdata.com>, public domain) and *geobr* (based on official IBGE shapefiles, public domain).

<https://doi.org/10.1371/journal.pntd.0014008.g001>

reference, resulting in 1,830,030 and 42,622 aligned bases, respectively (**Table C in S1 Text**). Genome coverage for SP17 was 100% and 80.1% for segments S and L, respectively. For SP19, the coverage was 57.7% for segment S and 14.2% for segment L. This is probably due to the lower viremia of SABV in SP19, as suggested by the lower fraction of SABV-specific reads in the metagenomic data (0.12% in SP17 versus 0.002% for SP19) and a longer period between symptom onset and sample collection (6 days for SP17 versus 9 days for SP19). Based on a 93.05% amino acid identity at the nucleoprotein (NP) level between SP17 and SP19 and the SABV reference strain SPH114202 (GenBank: NC_006317), our findings support that both genomes belong to SABV, a member of the *Brazilian mammarenavirus* species, consistent with the International Committee on Taxonomy of Viruses (ICTV) criterion for within-species classification in the *Arenavirus* genus [3]. Furthermore, SP17 and SP19 shared 98.19% amino acid identity at the NP, calculated across the overlapping aligned regions.

At the nucleotide level, the SABV genomes from SP17 and SP19 shared 73.06% identity in segment S and 84.53% in segment L. Compared to the SPH114202 strain (1990), SP17 showed 89.00% and 73.59% nucleotide identity in segments S and L, respectively (**Table D in S1 Text**), suggesting that SABV has been circulating in Brazil for decades and that its emergence in humans is not a recent evolutionary event.

Currently, SABV surveillance relies on the Bowen et al [12] protocol designed for the SPH114202 reference strain (1990) which is implemented in some of the Brazilian public health laboratories. We were able to identify two additional nucleotide mismatches to the 1696R primer in the SP17 genome compared to the SPH114202 strain (**Fig 2**), totaling 5 mismatches. Unfortunately, the genomic sequence of SP19 at the 1696R primer site was unavailable for comparison due to insufficient sequencing coverage in this region (**Table D in S1 Text**).

Because contemporary SABV genomes show mismatches with existing primers, multiple new assays were developed and evaluated. Based on the SPH114202 and newly generated genomes, we designed six new PCR primer sets and evaluated their performance in detecting contemporary SABV strains (**Tables D–F in S1 Text**). Primer sets 1 and 3 were selected for screening based on a combination of their amplification efficiency in preliminary tests and shorter amplicon length, which is advantageous for detecting potentially degraded RNA in field samples. These primer sets were subsequently applied to screen 29 (20 high exposure and nine symptomatic) healthcare workers who had either been in close proximity to SP17 or handled the patient's blood at *Hospital das Clínicas* and IMT, following the diagnostic framework previously described elsewhere [10]. Next, to detect SABV RNA in SP19, which had lower genome coverage and a reduced proportion of SABV-specific reads, we designed two nested PCR primer sets (**Tables G and H in S1 Text**). Our nested PCR assay successfully amplified SABV RNA from the SP19 sample, generating clear DNA bands on agarose gel, despite its low viremia (**Fig D in S1 Text**).

Maximum likelihood phylogenetic analysis of SABV together with 34 arenavirus species places the SP17, SP19 and SPH114202 strains within Clade B New World arenaviruses (**Figs 1 and 1B in S1 Text**). Both SP17 and SP19 strains clustered with SPH114202 with maximum bootstrap support. We found no evidence of interspecies reassortment or significant genome-wide recombination. These results were further corroborated by Bayesian phylogenetic analyses, which confirmed the SP17, SP19 and SPH114202 cluster with maximum posterior support (**Fig 3**). Using a phylogenetic model that takes into account time-dependency of evolutionary rates, we estimated the most recent common ancestor of SABV around 414 years ago (95% HPD: 142–750) for segment L (**Fig 3A**) and 1,081 years ago (95% HPD: 326–2,539) for segment S (**Fig 3B**). Conservatively, the lower 95% HPD bound suggests SABV has circulated for at least 142 years, with central estimates indicating several centuries. Segment L and S dating estimates overlap significantly, indicating

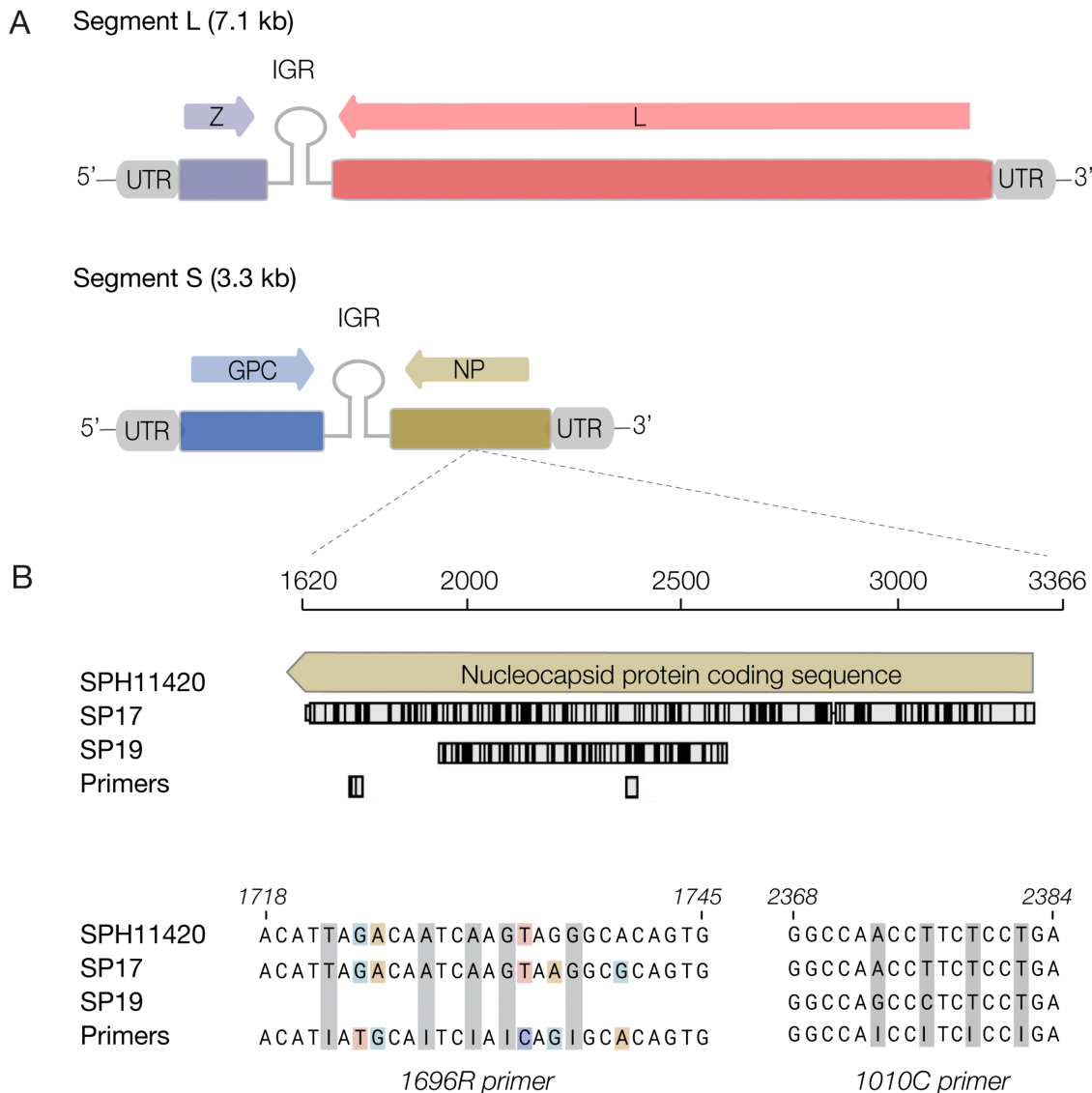


Fig 2. SABV genome structure and primer-binding site mismatches. (A) SABV genome structure, featuring an enveloped RNA with a single-stranded ambisense configuration. Segment L encodes the Z matrix protein and the L polymerase protein, while Segment S is responsible for encoding the GP and NP, separated by non-coding intergenic regions (IGR). Linear genomes are flanked by conserved 5' and 3' untranslated regions (UTRs) crucial for viral replication and translation processes [13]. (B) The upper section displays new strain sequences relative to the SABV reference, primer locations (Bowen et al. [12]), and variant positions in black. The lower part presents primer and SABV sequences at binding sites, with colours indicating primer mismatches and pale grey bars indicating inosine sites. Note: primers are shown in their published 5' → 3' orientation. Reverse primers, such as 1010C (5'-TCIGGIGAIGGITGGCC-3'), anneal to the genome by reverse complementarity and therefore may not appear as direct sequence matches when aligned against the 5'-3' representation of the negative-sense genome. Vs = versus, Z=matrix protein Z, L=RNA-dependent RNA polymerase protein L, GPC=glycoprotein precursor, NP=nucleocapsid protein, IGR=non-coding intergenic regions, UTR=untranslated regions.

<https://doi.org/10.1371/journal.pntd.0014008.g002>

robustness of the findings (Fig 3C). SABV shared an ancestral node with Chapare mammarenavirus, dating back to 0.015 million years ago (Mya) (95% HPD: 0.008–0.024 Mya) for segment L and 0.012 Mya (95% HPD: 0.006–0.021 Mya) for segment S. Clade B's emergence was estimated at approximately 0.17 Mya (95% HPD: 0.13–0.24 Mya) for segment L and 0.11 Mya (95% HPD: 0.06–0.16 Mya) for segment S. NWA and OWA originated around 3.77 Mya (95% HPD: 2.84–4.74 Mya) for segment L and 1.95 Mya (95% HPD: 1.40–2.61 Mya) for segment S.

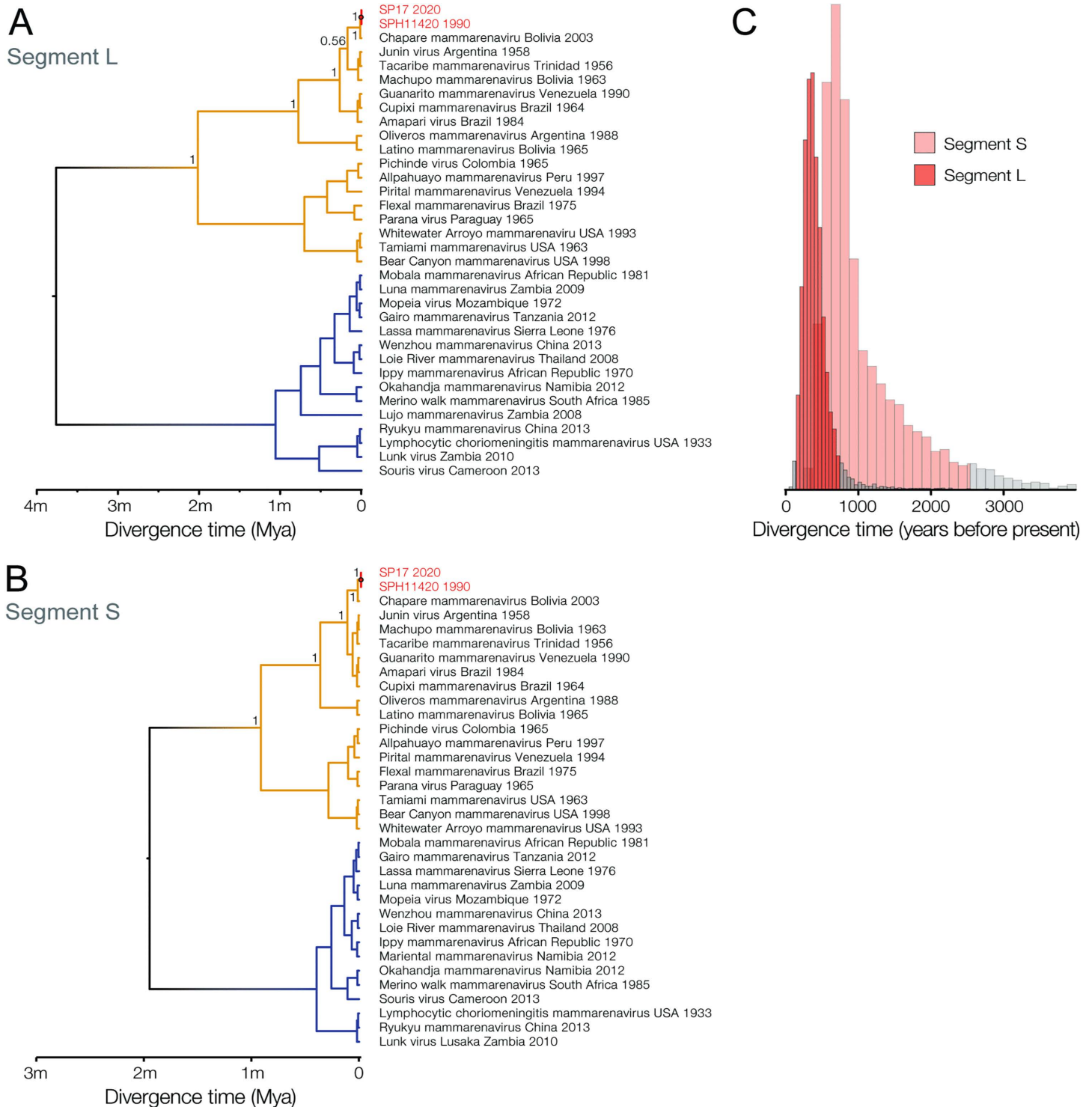


Fig 3. Time-calibrated phylogenies and estimated divergence times for SABV and other arenaviruses. We employ sequence data from SP17 with higher genome coverage for analysis. M= millions, Mya= million years. Inferences were performed for both S (A) and L (B) segments. (C) Posterior distributions for the dates of segments S and L most recent common ancestor.

<https://doi.org/10.1371/journal.pntd.0014008.g003>

The SABV GP is divided into stable signal peptide (SSP), GP1 attachment glycoprotein, and GP2 fusion glycoprotein subcomponents, which form a trimer of heterotrimers on the virion surface [14,15]. The GP complex of the new SABV strains share 94% amino acid sequence identity with SPH114202 [2,7] (Fig E in S1 Text). Our analysis shows that GP2 is more conserved than GP1, consistent with GP2 functional constraints described in other arenaviruses (Fig E in S1 Text). Specifically, receptor-binding GP1s have a nucleotide identity of 91% with 17 amino acid substitutions, while their GP2s (fusion glycoproteins) have an amino acid identity of 97%, with eight amino acid substitutions. Four amino acid substitutions – F5L, I36V, T51I, and R55K – are in the membrane-inserted SSP (Fig E in S1 Text). Similar variations have been reported in other arenaviral SSPs substitutions and unlikely to affect GP functionality [16].

We generated a homology model of SABV GP1 to examine whether the identified strains exhibit amino acid variation when compared with reference SPH114202 at the putative hTfR1 receptor-binding site (RBS) on the GP1 and thereby reflect different receptor tropism characteristics of the identified viruses (Fig 4A). All residues between the SABV GP1 reference and SP17 strains that align with hTfR1-interacting MACV GP1 residues are identical except N208D and R216T (corresponding to Thr216 and Ser224 on MACV GP1, respectively). When mapped onto the SABV GP1 (SP17) homology

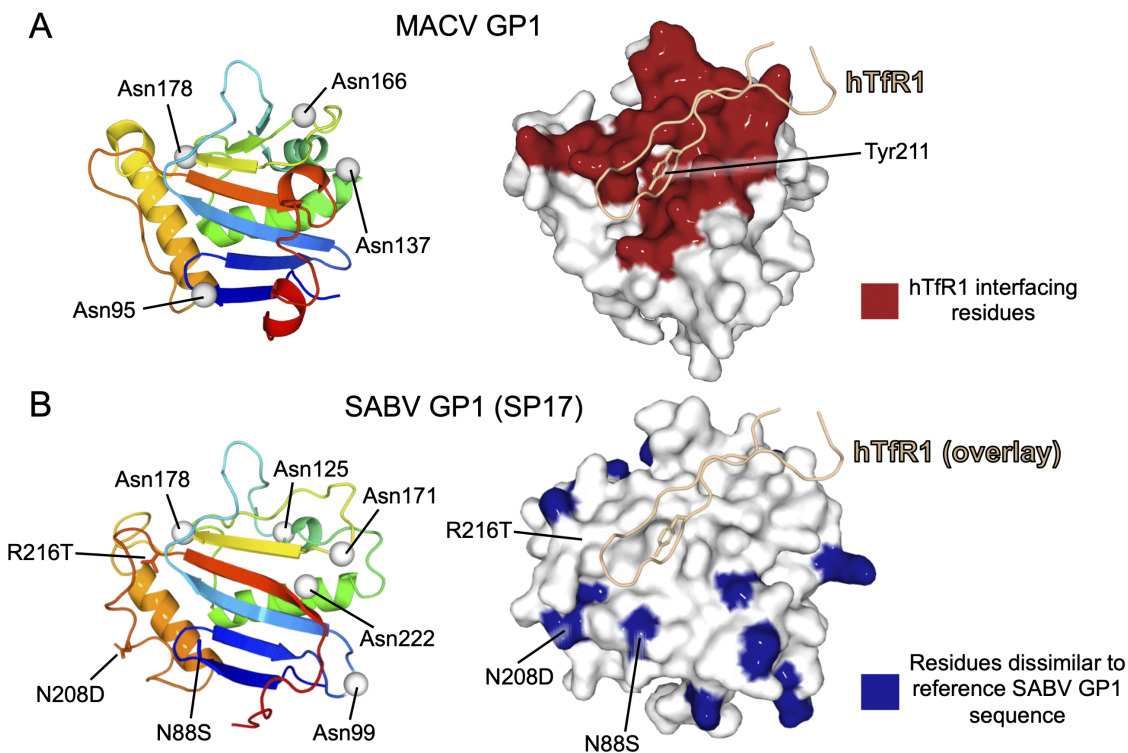


Fig 4. Structural analysis of SABV GP1 mutations. To facilitate comparative analysis of amino acid conservation between the SABV GP1 reference and contemporaneous strains we generated a homology model of SABV GP1 (SP17 strain) using AlphaFold3 [17]. (A) Mapping the TfR1-interaction region, an important determinant of zoonotic potential [18], on the surface of the GP1. (Left) Structure of MACV GP1 (PDB no. 3KAS) is shown as a rainbow-colored cartoon from blue (N-) to red (C-terminus) with N-linked glycosylation sites displayed as grey spheres. (Right) Structure of MACV GP1 is shown in white-colored surface representation, with hTfR1-binding residues colored red (determined with PDBEPIA [19]). A GP1-interacting chain of hTfR1 is shown as an orange-colored ribbon. The side chain of hTfR1-Tyr211 is shown as a stick. (B) Mapping amino acid conservation between SABV GP1 from the reference and contemporaneous strains. (Left) AlphaFold model of the new SABV strain GP1 (SP17), with putative N-linked glycosylation sites displayed as grey spheres. For fair comparison, only residues 80–234 are shown. Substitutions N88S, N208D, and R216T are displayed as sticks. (Right) The model is shown in surface representation and colored white. Amino acid residues dissimilar to the reference SABV GP1 are colored blue. A GP1-interacting chain of hTfR1 (PDB no. 3KAS) is overlaid onto the homology model of SABV GP1 to indicate region of the putative TfR1 binding site. [20].

<https://doi.org/10.1371/journal.pntd.0014008.g004>

model, D208 and T216 localize to the periphery of the putative RBS (Figs 4B, 4E and 4F in S1 Text). It seems plausible that both substitutions are unlikely to have a deleterious effect upon TfR1 binding, whereas N208D will result in minor electrostatic differences within the interface, and the formation of a smaller side-chain in R216T is unlikely to introduce steric hindrance to the interaction. These observations support the hypothesis that a selection pressure exists to maintain TfR1 recognition across SABV strains.

Our mapping analysis also indicates that other identified amino acid differences between SABV GP1 from the 1990 reference strain SPH114202 and the SP17 and SP19 SABV strains localize to putative solvent-accessible surfaces on the GP1 (Fig 4B). Given that there is currently no known functional role for these residues, which are placed outside of the putative TfR1-binding region, these substitutions may have arisen out of antigenic pressure. Similarly, the absence of an N-linked glycosylation sequon (N88S) in contemporaneous SABV strains suggests that this putative glycan may not be essential for native functionality or protein folding, and may have evolved in the original SABV sequence to shield protein-specific epitopes from the antibody immune response arising from infection, as has been previously observed for other arenaviruses [20].

Discussion

Genomic surveillance of recent fatal hemorrhagic cases identified two recent arenavirus infections in São Paulo, Brazil [10,11,21]. Our phylogenetic analysis showed that the viruses sequenced from SP17 and SP19 patient samples belong to the Sabiá virus species, a clade B New World arenavirus. The contemporaneous SABV SP17 strain exhibited variation at the nucleotide and amino acid level, evading detection by existing molecular assays, and our dating phylogenetic analyses indicate that SABV emerged at least over a century ago.

Epidemiological investigations suggest that all naturally occurring infections reported to date have had a history of exposure to rural areas across São Paulo state and resulted in fatal outcomes with diagnoses confirmed post-mortem. Although person-to-person and aerosol transmission are possible transmission routes [5,6], the two SABV cases reported here were linked to forest and rural activities, and no person-to-person transmission was detected among 29 close contacts of patient SP17. This suggests that the two SABV cases reported are most likely the result of distinct introductions from an unsampled reservoir, likely wild rodents [22]. Endemic areas for arenaviruses are expected to expand due to land use changes that increase contacts between reservoirs and humans [23], warranting increased surveillance of potential reservoir hosts. Moreover, three out of four naturally occurring SABV cases detected to date exhibited symptoms during the yellow fever transmission season in São Paulo where nearly half of the deaths attributed to hemorrhagic fever and neuroinvasive disease remain undiagnosed [24]. Prospective serological and molecular testing on patients from Sorocaba/Eldorado and Assis, as well as more broadly on YF-negative cases with hemorrhagic symptoms, could help provide a deeper understanding of SABV epidemiology, including its pathogenicity and transmissibility.

Our analyses of new SABV sequences uncovered substantial genetic diversity of SABV strains ($\geq 10\%$ nucleotide distance) in the NP gene. Synonymous substitutions and within-species diversity align with observations in other arenaviruses, such as Guanarito in Venezuela [25] and Lassa virus in West Africa [26]. Our study also provides a new perspective on an earlier report describing the SP17 strain, which suggested that this strain did not belong to a defined virus species, being characterized as SABV-like mammarenavirus [11]. While experimental verification is needed, mapping sequence conservation onto a model of SABV GP1 showed a high level of sequence conservation at the receptor-binding site and suggested that contemporaneous SABV strains share a TfR1-mediated host-cell entry pathway with the SABV reference strain. Thus, our molecular modeling suggests that contemporaneous SABV strains exhibit conserved species tropism and zoonotic potential characteristics.

Our phylogenetic analyses suggest that New World arenaviruses have circulated for millions of years [22] and show that SABV emerged approximately 1,081 and 414 years ago for the S and L segments, respectively. We show that contemporaneous SABV infections were not detected by existing molecular assays. In São Paulo, nearly half of the deaths

attributed to hemorrhagic fever and neuroinvasive disease remain undiagnosed [24]. Therefore, based on current data, it is likely that SABV infections have remained undetected and that the virus has circulated cryptically in São Paulo state, Brazil.

Four NWA virus species, Cupixi, Amapari, Flexal, and Sabiá, have been identified in Brazil. The potential limitations of current molecular diagnostic approaches identified here for SABV underscores the importance of untargeted metagenomic analyses to support outbreak investigations and to better understand NWA epidemiology, evolution and transmission dynamics. Reported SABV infections and other Clade B New World arenavirus infections are associated with high case fatality among detected cases [27,28]; however, this estimate may be inflated towards severe and fatal presentations. It is plausible that mild or subclinical SABV infections occur but remain undetected, as illustrated by the survival of both laboratory-associated cases. Although ribavirin has demonstrated benefit in Lassa fever and has been used empirically in other arenavirus infections [27], evidence supporting its efficacy against Sabiá virus infection is lacking, and clinical outcomes may reflect a spectrum of disease severity rather than antiviral effect. Experimental treatment options for SABV such as ribavirin, combined with the development of rapid point-of-care diagnostic methods, could be critical for patient management.

This study has several limitations. First, due to the unavailability of standardized positive controls (e.g., cloned RNA), we validated the PCR assays using SP17 and SP19 samples, which were genetically characterized through next-generation sequencing. Although this provided practical confirmation in the context of an urgent public health investigation, further validation of the assay's performance characteristics is required before broader surveillance implementation. Second, the low genome coverage of SP19 limited its use in this study. The absence of full-genome data constrained more comprehensive evolutionary and functional assessments. In addition, other confirmatory methods (e.g., Sanger sequencing) could not be performed due to limited remaining RNA. Virus isolation could not be performed because SABV is classified as a Risk Group 4 pathogen, and Brazil currently lacks the operational BSL-4 laboratory facilities required for such procedures. These constraints further limited the experimental confirmation of infectivity, receptor binding, and viral characteristics for SP19. Third, structural modeling of GP1 was based on limited sequence identity with other arenavirus glycoproteins and thus requires further structural and functional validation. Lastly, the ecological dynamics of SABV and its potential rodent reservoirs remain to be elucidated. Broader genomic, ecological, and serological surveillance efforts are essential to fill these gaps and better understand the risks posed by SABV and other New World arenaviruses in Brazil.

Routine, low-cost, and rapid pathogen-agnostic genomic surveillance is essential to track the circulation of new or divergent viruses and to improve molecular diagnostics. Analysis of YFV-negative genomic data identified this rare, highly lethal SABV as the causative agent. This study ruled out epidemiological links between recent cases and confirmed that close contacts were not infected. It also revealed SABV's evolutionary history and provided insights into its receptor-binding properties. These findings highlight SABV's potential for occasional spillover from hidden reservoirs, likely rodents, into human populations and underscore the urgent need to improve SABV diagnostic capacity in Brazil.

Materials and methods

Ethics statement

This study was approved by the *Comissão Nacional de Ética em Pesquisa* (CONEP), under protocol number CAAE 82673018.6.1001.0068. All patients signed an informed consent form to participate in the study and provided consent for the future use of their samples (BioBank informed consent).

Clinical-epidemiological context

The *Hospital das Clínicas* is the main tertiary care hospital associated with the Faculty of Medicine of the University of São Paulo. It offers specialized care to patients with complex medical conditions throughout Brazil. During the YF outbreak in Brazil between 2018 and 2019, the *Hospital das Clínicas* was essential in admitting severe YF cases. Concomitantly, a

randomized clinical trial - “SOFFA” study was carried out to assess the effectiveness of sofosbuvir in treating yellow fever [28]. Out of the 67 randomized patients, eight individuals tested negative for YFV RNA in their serum and urine samples, as well as other causes of haemorrhagic fever (HF) (such as dengue virus, hepatovirus A, B, C, Delta, and E, HIV, CMV, EBV, HSV, malaria, leptospirosis, toxoplasmosis, and bacterial septic shock).

Viral metagenomics

Serum samples from the eight YFV-negative patients were centrifuged at 10,000g for 5 minutes. Then, viral RNA was extracted from the sample supernatant using QIAamp RNA viral extraction kit (Qiagen, Valencia, CA, USA) in a Biosafety Level-3 facility (BSL-3) at the *Instituto de Medicina Tropical* (IMT). Next, extracted RNA was subjected to untargeted metagenomic sequencing using the SMART-9N protocol [29]. Libraries were prepared using the SQK-LSK109 ligation sequencing kit and sequenced on the MinION platform (Oxford Nanopore Technologies, ONT, UK). Additional protocol details are provided in [S1 Text](#).

Each sample was subjected to two full technical replicates before sequencing on separate and previously unused R9.4.1 flow cells (ONT, UK), with a negative control added in each reaction. Raw files were demultiplexed using Guppy in High Accuracy mode (ONT, UK). Fastq files were taxonomically classified using the standard Kraken2 database [30], and the results were visualized using the pavian tool [31]. Identified reads were mapped against the SABV reference genome (S and L segments, accession numbers AY358026 and U41071) with minimap2 version 2.17-r941 with parameters ‘-a’ (SAM output) and ‘-x map-ont’ (nanopore mapping mode), ‘-w1’ and ‘-k11’ (increased sensitivity) [32].

Genome statistics were assessed with Samtools [33] and the Tablet viewer [34]. To recover consensus sequences from the S and L genomic fragments, we called variants detected with Medaka [35] for regions of the genome covered with at least 5 reads. Consensus sequences were subjected to BLAST searches against the NCBI viral protein and nucleotide sequence databases, and the percent identity were determined between the new SABV genomes and available SABV reference sequences [36]. The sequences, along with available information, were immediately shared with the São Paulo Secretary of Health and included in a report to support further research and public health initiatives [21]. The genome sequences of sample SP17 have been submitted to GenBank under accession numbers PX308626 and PX308627, while the partial genome sequences of sample SP19 are available at https://github.com/CADDE-CENTRE/SABV_Brazil.

Phylogenetic analyses

To investigate SABV diversity within the *Arenaviridae* family, phylogenetic analyses were conducted for the L and S genome segments. Reference datasets from the International Committee on Taxonomy of Viruses (ICTV) included a nucleotide (nt) and amino acid (aa) genome sequence per viral species from Clades A, B, and C of the New World arenavirus (NWA) and the Old World arenavirus (OWA) complexes (n = 36, dataset A - https://github.com/CADDE-CENTRE/SABV_Brazil). Phylogenies were estimated using GTR nucleotide substitution model and a blosum62 aa exchange matrix for nt and aa-based datasets, respectively, in PhyML [37] using Seaview [38]. Branch support was assessed using an approximate likelihood ratio test [39].

To confirm the phylogenetic placement of the newly generated SABV sequences, we downloaded 6,143 coding sequences of the Mammarenavirus genus from the NCBI Virus platform (<https://ncbi.nlm.nih.gov/labs/virus/vssi/#/virus>). We filtered sequences for the NP, GP, Z, and RdRp ORFs based on length thresholds, set at 70% of the corresponding protein length in the Lassa mammarenavirus (LASV) RefSeq genome (segment S: NC_004296, segment L: NC_004297). Sequences lacking collection dates or containing stop codons were removed, yielding reference datasets of 862 (NP), 891 (GP), 708 (Z), and 755 (RdRp) nucleotide sequences, respectively (Dataset B, available at https://github.com/CADDE-CENTRE/SABV_Brazil). The final datasets were combined with the newly generated SABV sequences, and phylogenetic trees were inferred using IQ-TREE v2 [40] with the GTR+F+I+G4 nucleotide substitution model and 1,000 maximum-likelihood bootstrap replicates.

To infer the evolutionary history of SABV, dated phylogenetic trees were estimated from the nucleotide dataset A in BEAST v1.10.4 [41] using the BEAGLE v3 library [41]. We applied a time-dependent rate (TDR) model [42] on the deep SABV history, with an informative prior on the modern rate (a normal distribution with mean -6.7 and standard deviation of 0.25 on the log rate) based on published LASV virus evolutionary rates [43] and an informative prior on the TDR decline rate (normal with mean -6.21 and standard deviation of 0.035) based on previous TDR estimates [42]. Furthermore, we specified a GTR+ Γ substitution model and a Yule tree prior across the S and L segments. Posterior distributions of the time of the most recent common ancestor of SABV strains from the S and L dated phylogenetic distributions were obtained using TreeStat v.10.4 [41]. Potential recombination events of the NWA were investigated from dataset A, using all available methods in RDP4 [44].

Homology modeling and structure-based mapping analysis of SABV GP1

Amino acid sequences of SABV envelope glycoprotein GP1 (F59–S250; Sample 17 (SP17)) were submitted to AlphaFold 3 server [17], with no additional customisation. SABV GP1 model with a predicted template modelling (pTM) score of 0.73 was chosen and used for structural-based mapping analysis, performed in PyMol - <https://pymol.org>. Putative asparagine (N)-linked glycosylation sites were predicted using NetNGlyc [45].

Development of a PCR and sensitive nested-cPCR assays

To investigate whether primer sets from current assays would detect the new SABV strains, we evaluated the performance of six new RT-PCR primer pairs (**Table E in S1 Text**). These primers were designed based on the 1990 reference genome of SABV and sequences newly obtained in this study. We selected conserved regions within the NP gene to ensure assay sensitivity and specificity. Degenerate bases were included in regions with sequence variation to improve compatibility with circulating strains. Briefly, the extracted RNA from SP17 and SP19 patient samples, and a negative control was converted into cDNA using Random Hexamers (Thermo Fisher Scientific, USA) and ProtoScript II First Strand cDNA Synthesis Kit (New England Biolabs, UK) following manufacturer's instructions. The RT-PCR reactions were prepared using 5 μ l Q5 reaction buffer (New England BioLabs, USA), 0.5 μ l 10 μ M dNTP, 1.25 μ l of the forward primer, and 1.25 μ l of the reverse primer (variable, according to the primer set, 10 μ M), 14.25 μ l Nuclease-free water (NFW), 0.25 μ l Q5 DNA polymerase, and 2.5 μ l of cDNA (**Tables F and G in S1 Text**). PCR cycling conditions were: 98°C for 45 sec, followed by 30 cycles of 98°C for 15 sec, T_m variable for 15 sec, and 65°C for 5 min and a final step of 65°C for 10 min, changing the T_m according to the set of primers used (**Tables E–G in S1 Text**). A gel was prepared using the E-gel EX Agarose 2% (Thermo Fisher Scientific, USA). To load the PCR products, 5 μ l of the products were mixed with 15 μ l of NFW and 20 μ l of E-gel Sizing DNA leader (Thermo Fisher Scientific, USA) was used in the first band. The gel was placed into the E-gel equipment (Thermo Fisher Scientific, USA) and the run was performed until the bands were distinguishable by transillumination.

We next designed two nested-cPCR primer sets in an attempt to detect low viremia cases (**Table H in S1 Text**). The nested-cPCR reactions were prepared using 5 μ l Q5 reaction buffer (New England BioLabs, USA), 0.5 μ l 10 μ M dNTP, 1 μ l of the forward primer, and 1 μ l of the reverse primer (variable, according to the primer set, 10 μ M), 14.75 μ l Nuclease-free water (NFW), 0.25 μ l Q5 DNA polymerase, and 2.5 μ l of cDNA from SP19 (**Tables H–J in S1 Text**). PCR cycling conditions were: 98°C for 45 sec, followed by 30 cycles of 98°C for 15 sec, T_m variable for 15 sec, and 65°C for 5 min and a final step of 65°C for 10 min. This reaction was done twice, once for the outer PCR reaction, and then for the inner PCR reaction, changing the T_m according to the set of primers used (**Tables H–J in S1 Text**). 5 μ l of the products were mixed with 15 μ l of NFW and 20 μ l of E-gel Sizing DNA leader were loaded to the bands visualization (Thermo Fisher Scientific, USA). The gel was placed into the E-gel equipment (Thermo Fisher Scientific, USA) and the run was performed until the bands were distinguishable by transillumination.

Following validation using SABV-positive clinical samples, selected primer sets were applied to screen 29 high-risk healthcare worker contacts, as part of the contact investigation previously described by Nastri et al. [10]. Detailed

laboratory protocols used in this study have been deposited in protocols.io and are publicly available at: <https://dx.doi.org/10.17504/protocols.io.4r3l2z293l1y/v1>.

Supporting information

S1 Text. Fig A. Pavian plots of Kraken2 database matches. Pavian diagrams of Kraken 2 report results obtained from SP4. The width of the flow is proportional to the number of reads. The number above each node is the number of k-mer hits. A rank code, indicating domain (D), phylum (P), family (F), genus (G), or species (S) was used. **Fig B.** Phylogenetic trees for Z, L, GP and NP SABV proteins. Maximum Likelihood phylogenetic trees of arenavirus species based on nucleotide (nt) alignments of Dataset B (available at <https://github.com/CADDE-CENTRE/Sabia>). The trees show two groups of arenaviruses, the Old World arenaviruses and the New World arenaviruses. The scale bar represents the units of nucleotide substitutions per site. SABV strains are distinctly highlighted with a box colored in a green gradient. The green tips indicate the patient's samples SP17 and SP19 described in this study. GPC = glycoprotein complex, Z = zinc-finger-like protein NP = nucleoprotein, RdRP = RNA-dependent RNA polymerase also known as L protein. **Fig C.** RT-PCR validation for different SABV primer sets. Left panel: the bands highlighted by a circle refer to the primer sets 1 (350 bp), 2 (200 bp), and 3 (121 bp) respectively; Right panel: no band was detected using primer set 4 and the bands highlighted by a circle refer to the primer sets 5 (723 bp) and 6 (200 bp), respectively. (NC: Negative control; bp: base pairs; Sample 17 = SP17; Sample 19 = SP19). **Fig D.** Nested cPCR validation for SABV. The bands highlighted by a circle refer to primer sets 1 (200 bp) and 2 (112 bp). (NC: Negative control; bp: base pairs; Sample 19 = SP19). **Fig E.** Sequence alignments of SABV GP1. (Top) An amino acid sequence alignment of the SABV GP complex from SP17 and reference SPH114202 strains. A sequence consensus line is displayed below the alignment. A highly conserved residue appears in high-consensus red color and as an uppercase letter in the consensus line while a weakly conserved residue is colored blue and as a lowercase letter. A position with no conserved residue is represented by a dot in the consensus line. Symbol! represents I and V, \$ for L and M, % for F and Y, # for N, D, Q, and E. (Bottom) An amino acid sequence alignment of MACV GP1, SABV SPH114202 GP1, and SABV GP1 from SP17. The sequence consensus and color scheme are as the top panel. Both alignments were plotted using MultAlin (13). **Fig F.** Analysis of MACV GP1 residues that interact with hTfR1. The structure of MACV GP1 in complex with hTfR1 (PDB no. 3KAS) was submitted to PDBEPIISA to analyze the interaction interface. The calculated results show accessible/buried surface area (ASA/BSA; Å²), solvation energy effect (ΔiG; kcal/mol), and intermolecular interactions (HSDC stands for hydrogen bond, salt bridge, disulphide bond and covalent link). Buried area percentage is represented by vertical bars (one bar per 10%). Residues on light blue background are solvent accessible residues while those on a darker blue shade are inaccessible residues. Only MACV GP1 ("Structure 2") residues 86–241 are shown here. **Table A.** Epidemiological characteristics of eight YFV-negative cases. ID = identity; Patients were also previously tested negative for the viruses dengue, hepatovirus A, B, C Delta, and E, HIV, CMV, EBV, and HSV, as well as malaria, leptospirosis, toxoplasmosis, and bacterial septic shock. **Table B.** Epidemiological characteristics of all reported SABV cases. YFV = Yellow fever virus; USA = United States of America; NA = natural acquired. Cases from Dec 2019 were most likely infected in Assis (420 km from Cotia) and Sorocaba/Eldorado (65/210 km from Cotia). Distance between Assis and Sorocaba or Eldorado is 356/430 km. All patients had a fatal clinical outcome. No genomic report exists of strain Case 2 (SPH185338). **Table C.** Sequencing and alignment statistics for the SABV and HAV cases analyzed in this report, including sequencing coverage statistics. **Table D.** Nucleotide and amino acid identity among SABV strains. **Table E.** SABV primer sets used for validation of an RT-PCR assay. Primers were designed based on the 1990 reference sequence of SABV and newly generated genomes from this study. Conserved regions of the nucleoprotein (NP) gene were selected as the amplification targets due to their diagnostic relevance and sequence conservation across SABV strains. **Table F.** Reagents and conditions for the SABV RT-PCR protocol. **Table G.** PCR cycling conditions for the SABV RT-PCR detection protocol. **Table H.** Primer sequences used in the nested-cPCR reactions. **Table I.** Nested cPCR reagents and conditions for the SABV detection protocols. Note that the reaction is done twice, once for the outer PCR reaction, and then

for the inner PCR reaction. **Table J.** PCR cycling conditions for the SABV nested cPCR detection protocol. Note that the reaction is done twice, once for the outer PCR reaction, and then for the inner PCR reaction.

(DOCX)

Author contributions

Conceptualization: Ingra M. Claro, Camila M. Romano, Sarah C. Hill, Darlan S. Candido, Mariana S. Ramundo, Weng M. Ng, Thomas A. Bowden, Anna S. Levin, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Data curation: Ingra M. Claro, Erika R. Manuli, Camila A. M. da Silva, Thais M. Coletti, Philippe Lemey, Ana Catharina Nastri, Luciana Vilas Boas Casadio, Amaro Nunes Duarte-Neto, Joshua Quick, Camila M. Romano, Sarah C. Hill, Filipe R. R. Moreira, Ian N. Valença, Jaqueline G. de Jesus, Flavia C. S. Sales, Weng M. Ng, Thomas A. Bowden, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Formal analysis: Ingra M. Claro, Philippe Lemey, Joshua Quick, Camila M. Romano, Sarah C. Hill, Carlos A. Prete, Darlan S. Candido, Filipe R. R. Moreira, Mariana S. Ramundo, Ian N. Valença, Jaqueline G. de Jesus, Weng M. Ng, Thomas A. Bowden, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Funding acquisition: Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Investigation: Ingra M. Claro, Erika R. Manuli, Camila A. M. da Silva, Thais M. Coletti, Philippe Lemey, Ana Catharina Nastri, Luciana Vilas Boas Casadio, Joshua Quick, Camila M. Romano, Charles Whittaker, Sarah C. Hill, Darlan S. Candido, Filipe R. R. Moreira, Ian N. Valença, Jaqueline G. de Jesus, Flavia C. S. Sales, Mariana S. Cunha, Juliana M. Guerra, Maria Cassia Mendes-Correa, Tania R. Tozetto-Mendoza, Marcilio Jorge Fumagalli, Yeh-Li Ho, Peter Simmonds, Weng M. Ng, Thomas A. Bowden, William M. de Souza, Oliver G. Pybus, Anna S. Levin, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Methodology: Ingra M. Claro, Erika R. Manuli, Camila A. M. da Silva, Thais M. Coletti, Philippe Lemey, Joshua Quick, Weng M. Ng, Thomas A. Bowden, William M. de Souza, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Project administration: Ingra M. Claro, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Resources: Ingra M. Claro, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Software: Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Supervision: Ingra M. Claro, Camila M. Romano, Anna S. Levin, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Validation: Ingra M. Claro, Philippe Lemey, Joshua Quick, Camila M. Romano, Sarah C. Hill, Carlos A. Prete, Maria Cassia Mendes-Correa, Weng M. Ng, Thomas A. Bowden, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Visualization: Ingra M. Claro, Philippe Lemey, Amaro Nunes Duarte-Neto, Joshua Quick, Camila M. Romano, Charles Whittaker, Sarah C. Hill, Carlos A. Prete, Darlan S. Candido, Filipe R. R. Moreira, Mariana S. Ramundo, Ian N. Valença, Jaqueline G. de Jesus, Flavia C. S. Sales, Mariana S. Cunha, Juliana M. Guerra, Maria Cassia Mendes-Correa, Tania R. Tozetto-Mendoza, Marcilio Jorge Fumagalli, Yeh-Li Ho, Peter Simmonds, Weng M. Ng, Thomas A. Bowden, William M. de Souza, Oliver G. Pybus, Anna S. Levin, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Writing – original draft: Ingra M. Claro, Erika R. Manuli, Camila A. M. da Silva, Thais M. Coletti, Philippe Lemey, Ana Catharina Nastri, Luciana Vilas Boas Casadio, Amaro Nunes Duarte-Neto, Joshua Quick, Camila M. Romano, Charles Whittaker, Sarah C. Hill, Carlos A. Prete, Darlan S. Candido, Filipe R. R. Moreira, Mariana S. Ramundo, Ian N. Valença, Jaqueline G. de Jesus, Flavia C. S. Sales, Mariana S. Cunha, Juliana M. Guerra, Maria Cassia Mendes-Correa, Tania R. Tozetto-Mendoza, Marcilio Jorge Fumagalli, Yeh-Li Ho, Peter Simmonds, Weng M. Ng, Thomas A.

Bowden, William M. de Souza, Oliver G. Pybus, Anna S. Levin, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

Writing – review & editing: Ingra M. Claro, Erika R. Manuli, Camila A. M. da Silva, Thais M. Coletti, Philippe Lemey, Ana Catharina Nastri, Luciana Vilas Boas Casadio, Amaro Nunes Duarte-Neto, Joshua Quick, Camila M. Romano, Charles Whittaker, Sarah C. Hill, Carlos A. Prete, Darlan S. Candido, Filipe R. R. Moreira, Mariana S. Ramundo, Ian N. Valença, Jaqueline G. de Jesus, Flavia C. S. Sales, Mariana S. Cunha, Juliana M. Guerra, Maria Cassia Mendes-Correa, Tania R. Tozetto-Mendoza, Marcilio Jorge Fumagalli, Yeh-Li Ho, Peter Simmonds, Weng M. Ng, Thomas A. Bowden, William M. de Souza, Oliver G. Pybus, Anna S. Levin, Nicholas Loman, Ester C. Sabino, Nuno Rodrigues Faria.

References

1. Kuhn JH, Abe J, Adkins S, Alkhovsky SV, Avšič-Županc T, Ayllón MA, et al. Annual (2023) taxonomic update of RNA-directed RNA polymerase-encoding negative-sense RNA viruses (realm Riboviria: kingdom Orthornavirae: phylum Negarnaviricota). *J Gen Virol.* 2023;104(8):001864. <https://doi.org/10.1099/jgv.0.001864> PMID: 37622664
2. Lisieux T, Coimbra M, Nassar ES, Burattini MN, de Souza LT, Ferreira I, et al. New arenavirus isolated in Brazil. *Lancet.* 1994;343(8894):391–2. [https://doi.org/10.1016/s0140-6736\(94\)91226-2](https://doi.org/10.1016/s0140-6736(94)91226-2) PMID: 7905555
3. Radoshitzky SR, Buchmeier MJ, Charrel RN, Clegg JCS, Gonzalez J-PJ, Günther S, et al. ICTV virus taxonomy profile: arenaviridae. *J Gen Virol.* 2019;100(8):1200–1. <https://doi.org/10.1099/jgv.0.001280> PMID: 31192784
4. Coimbra TLM, Santos RN, Ferreira IB, Fialho DM, Mello ES, Ferreira LMHL, et al. Arenavirus: A fatal outcome. *VR&R.* 2001;6(1). <https://doi.org/10.17525/vrrjournal.v6i1.187>
5. Gandsman EJ, Aaslestad HG, Ouimet TC, Rupp WD. Sabia virus incident at Yale University. *Am Ind Hyg Assoc J.* 1997;58(1):51–3. <https://doi.org/10.1080/15428119791013080> PMID: 9018837
6. Vasconcelos PF, Travassos da Rosa AP, Rodrigues SG, Tesh R, Travassos da Rosa JF, Travassos da Rosa ES. Laboratory-acquired human infection with SP H 114202 virus (Arenavirus: Arenaviridae family): clinical and laboratory aspects. *Rev Inst Med Trop Sao Paulo.* 1993;35(6):521–5. <https://doi.org/10.1590/s0036-46651993000600008> PMID: 7997756
7. Gonzalez JP, Bowen MD, Nichol ST, Rico-Hesse R. Genetic characterization and phylogeny of Sabiá virus, an emergent pathogen in Brazil. *Virology.* 1996;221(2):318–24. <https://doi.org/10.1006/viro.1996.0381> PMID: 8661442
8. Choe H, Jemielity S, Abraham J, Radoshitzky SR, Farzan M. Transferrin receptor 1 in the zoonosis and pathogenesis of New World hemorrhagic fever arenaviruses. *Curr Opin Microbiol.* 2011;14(4):476–82. <https://doi.org/10.1016/j.mib.2011.07.014> PMID: 21807555
9. Pfau CJ. Arenaviruses. In: Baron S, editor. *Medical Microbiology.* 4th ed. Galveston (TX). 1996.
10. Nastri AC, Duarte-Neto AN, Casadio LVB, Souza WM de, Claro IM, Manuli ER, et al. Understanding Sabiá virus infections (Brazilian mammarenavirus). *Travel Med Infect Dis.* 2022;48:102351. <https://doi.org/10.1016/j.tmaid.2022.102351> PMID: 35537676
11. de Mello Malta F, Amgarten D, Nastri AC de SS, Ho Y-L, Boas Casadio LV, Basqueira M, et al. Sabiá Virus-Like Mammarenavirus in Patient with Fatal Hemorrhagic Fever, Brazil, 2020. *Emerg Infect Dis.* 2020;26(6):1332–4. <https://doi.org/10.3201/eid2606.200099> PMID: 32441627
12. Bowen MD, Peters CJ, Nichol ST. The phylogeny of New World (Tacaribe complex) arenaviruses. *Virology.* 1996;219(1):285–90. <https://doi.org/10.1006/viro.1996.0248> PMID: 8623541
13. McLay L, Ansari A, Liang Y, Ly H. Targeting virulence mechanisms for the prevention and therapy of arenaviral hemorrhagic fever. *Antiviral Res.* 2013;97(2):81–92. <https://doi.org/10.1016/j.antiviral.2012.12.003> PMID: 23261843
14. Hulswit RJG, Paesen GC, Bowden TA, Shi X. Recent advances in bunyavirus glycoprotein research: precursor processing, receptor binding and structure. *Viruses.* 2021;13(2):353. <https://doi.org/10.3390/v13020353> PMID: 33672327
15. Paesen GC, Ng WM, Kimuda S, Sutton G, Doores KJ, Bowden TA. Structure and stabilization of the antigenic glycoprotein building blocks of the New World mammarenavirus spike complex. *mBio.* 2025;16(7):e0107625. <https://doi.org/10.1128/mbio.01076-25> PMID: 40511941
16. Nunberg JH, York J. The curious case of arenavirus entry, and its inhibition. *Viruses.* 2012;4(1):83–101. <https://doi.org/10.3390/v4010083> PMID: 22355453
17. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature.* 2024;630(8016):493–500. <https://doi.org/10.1038/s41586-024-07487-w> PMID: 38718835
18. Abraham J, Corbett KD, Farzan M, Choe H, Harrison SC. Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. *Nat Struct Mol Biol.* 2010;17(4):438–44. <https://doi.org/10.1038/nsmb.1772> PMID: 20208545
19. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol.* 2007;372(3):774–97. <https://doi.org/10.1016/j.jmb.2007.05.022> PMID: 17681537

20. Sommerstein R, Flatz L, Remy MM, Malinge P, Magistrelli G, Fischer N, et al. Arenavirus glycan shield promotes neutralizing antibody evasion and protracted infection. *PLoS Pathog.* 2015;11(11):e1005276. <https://doi.org/10.1371/journal.ppat.1005276> PMID: 26587982
21. Cadde-centre. SABV report. Cadde-centre; 2020. https://github.com/CADDE-CENTRE/Miscellaneous/blob/master/SABV_Report.pdf
22. Sarute N, Ross SR. New world arenavirus biology. *Annu Rev Virol.* 2017;4(1):141–58. <https://doi.org/10.1146/annurev-virology-101416-042001> PMID: 28645238
23. Clegg JC. Influence of climate change on the incidence and impact of arenavirus diseases: a speculative assessment. *Clin Microbiol Infect.* 2009;15(6):504–9. <https://doi.org/10.1111/j.1469-0691.2009.02847.x> PMID: 19604274
24. Araújo LJ de, Gonzalez LL, Buss LF, Guerra JM, Gomez DS, Ferreira CS da S, et al. Surveillance of hemorrhagic fever and/or neuroinvasive disease: challenges of diagnosis. *Rev Saude Publica.* 2021;55:41. <https://doi.org/10.11606/s1518-8787.2021055003068> PMID: 34190890
25. Fulhorst CF, Cajimat MNB, Milazzo ML, Paredes H, de Manzione NMC, Salas RA, et al. Genetic diversity between and within the arenavirus species indigenous to western Venezuela. *Virology.* 2008;378(2):205–13. <https://doi.org/10.1016/j.virol.2008.05.014> PMID: 18586298
26. Leski TA, Stockelman MG, Moses LM, Park M, Stenger DA, Ansumana R, et al. Sequence variability and geographic distribution of Lassa virus, Sierra Leone. *Emerg Infect Dis.* 2015;21(4):609–18. <https://doi.org/10.3201/eid2104.141469> PMID: 25811712
27. McCormick JB, King IJ, Webb PA, Scribner CL, Craven RB, Johnson KM, et al. Lassa fever. Effective therapy with ribavirin. *N Engl J Med.* 1986;314(1):20–6. <https://doi.org/10.1056/NEJM198601023140104> PMID: 3940312
28. Figueiredo-Mello C, Casadio LVB, Avelino-Silva VI, Yeh-Li H, Sztajn bok J, Joelsons D, et al. Efficacy of sofosbuvir as treatment for yellow fever: protocol for a randomised controlled trial in Brazil (SOFFA study). *BMJ Open.* 2019;9(11):e027207. <https://doi.org/10.1136/bmjopen-2018-027207> PMID: 31772079
29. Claro IM, Ramundo MS, Coletti TM, da Silva CAM, Valenca IN, Candido DS, et al. Rapid viral metagenomics using SMART-9N amplification and nanopore sequencing. *Wellcome Open Res.* 2023;6:241. <https://doi.org/10.12688/wellcomeopenres.17170.2> PMID: 37224315
30. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* 2019;20(1):257. <https://doi.org/10.1186/s13059-019-1891-0> PMID: 31779668
31. Breitwieser FP, Salzberg SL. Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics.* 2020;36(4):1303–4. <https://doi.org/10.1093/bioinformatics/btz715> PMID: 31553437
32. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics.* 2018;34(18):3094–100. <https://doi.org/10.1093/bioinformatics/bty191> PMID: 29750242
33. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
34. Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, et al. Tablet—next generation sequence assembly visualization. *Bioinformatics.* 2010;26(3):401–2. <https://doi.org/10.1093/bioinformatics/btp666> PMID: 19965881
35. Technologies ON. Medaka: sequence correction provided by ONT. Oxford Nanopore Technologies; 2024.
36. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: 2231712
37. Guindon S, Delsuc F, Dufayard J-F, Gascuel O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol.* 2009;537:113–37. https://doi.org/10.1007/978-1-59745-251-9_6 PMID: 19378142
38. Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 2010;27(2):221–4. <https://doi.org/10.1093/molbev/msp259> PMID: 19854763
39. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol.* 2006;55(4):539–52. <https://doi.org/10.1080/10635150600755453> PMID: 16785212
40. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the Genomic Era. *Mol Biol Evol.* 2020;37(5):1530–4. <https://doi.org/10.1093/molbev/msaa015> PMID: 32011700
41. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 2012;29(8):1969–73. <https://doi.org/10.1093/molbev/mss075> PMID: 22367748
42. Membrebe JV, Suchard MA, Rambaut A, Baele G, Lemey P. Bayesian Inference of evolutionary histories under time-dependent substitution rates. *Mol Biol Evol.* 2019;36(8):1793–803. <https://doi.org/10.1093/molbev/msz094> PMID: 31004175
43. Klitting R, Kafetzopoulou LE, Thiery W, Dudas G, Gryseels S, Kotamarthi A, et al. Predicting the evolution of the Lassa virus endemic area and population at risk over the next decades. *Nat Commun.* 2022;13(1):5596. <https://doi.org/10.1038/s41467-022-33112-3> PMID: 36167835
44. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* 2015;1(1):vev003. <https://doi.org/10.1093/ve/vev003> PMID: 27774277
45. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput.* 2002;310–22. PMID: 11928486