Virus Evolution, 2025, **11(1)**, veae102
DOI: https://doi.org/10.1093/ve/veae102
Advance Access Publication Date: 23 November 2024
Research Article

Comprehensive molecular epidemiology of influenza viruses in Brazil: insights from a nationwide analysis

Isabela Carvalho Brcko[1,2], Vinicius Carius de Souza[1,2], Gabriela Ribeiro[1,2], Alex Ranieri Jeronimo Lima[1,2], Antonio Jorge Martins[1,2], Claudia Renata dos Santos Barros[1,2], Eneas de Carvalho[2], James Siqueira Pereira[1,2], Loyze Paola Oliveira de Lima[1,2], Vincent Louis Viala[1,2], Simone Kashima[1,3], Debora Glenda Lima de La Roque[3], Elaine Vieira Santos[3], Evandra Strazza Rodrigues[3], Juliana Almeida Nunes[4], Leandro Spalato Torres[4], Luiz Artur Vieira Caldeira[4], Melissa Palmieri[4], Caio Genovez Medina[5], Raphael Augusto de Arruda[5], Renata Beividas Lopes[5], Geraldo Reple Sobrinho[6], Daniel Macedo de Melo Jorge[7], Eurico Arruda[7], Eladja Christina Bezerra da Silva Mendes[8], Hazerral de Oliveira Santos[8], Arabela Leal e Silva de Mello[9], Felicidade Mota Pereira[9], Marcela Kelly Astete Gómez[9], Vanessa Brandão Nardy[9], Brenno Henrique[10], Lucas Luiz Vieira[10], Mariana Matos Roll[10], Elaine Cristina de Oliveira[11], Júlia Deffune Profeta Cidin Almeida[11], Stephanni Figueiredo da Silva[11], Gleissy Adriane Lima Borges[12], Katia Cristina de Lima Furtado[12], Patricia Miriam Sayuri Sato Barros da Costa[12], Shirley Moreira da Silva Chagas[12], Esper G. Kallás[2,13], Daniel Larh[14], Marta Giovanetti[15,16,17], Svetoslav Nanev Slavov[1,2,*], Sandra Coccuzzo Sampaio[1,2,*], Maria Carolina Elias[1,2,*]

[1]Center for Viral Surveillance and Serological Assessment (CeVIVAS), Instituto Butantan, Avenida Vital Brasil, 1500, Butantã, São Paulo, São Paulo 05503-900, Brazil
[2]Instituto Butantan, Avenida Vital Brasil, 1500, Butantã, São Paulo, São Paulo 05503-900, Brazil
[3]Hemocentro Ribeirão Preto, Rua Tenente Catão Roxo, 2501, Ribeirão Preto, São Paulo 14051-140, Brazil
[4]Coordenadoria de Vigilância em Saúde (COVISA), Secretaria Municipal de Saúde (SMS/SP), Prefeitura São Paulo, Rua Siqueira Campos, 176, São Paulo, São Paulo 01509-020, Brazil
[5]Departamento de Atenção Hospitalar de São Bernardo do Campo, Prefeitura São Bernardo do Campo, Rua João Pessoa, 59, São Bernardo do Campo, São Paulo 09715-000, Brazil
[6]Secretaria de Saúde de São Bernardo do Campo, Prefeitura São Bernardo do Campo, Rua João Pessoa, 59, São Bernardo do Campo, São Paulo 09715-000, Brazil
[7]Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Avenida dos Bandeirantes, 3900, Ribeirão Preto, São Paulo 14048-900, Brazil
[8]Laboratório Central de Saúde Pública do Estado de Alagoas (LACEN-AL), Rua Doutor Ernesto Gomes Maranhão, 1773, Maceió, Alagoas 57036-860, Brazil
[9]Laboratório Central de Saúde Pública do Estado da Bahia (LACEN-BA), 4ª Avenida, 400, Salvador, Bahia 41745-900, Brazil
[10]Laboratório Central de Saúde Pública do Distrito Federal (LACEN-DF), Lotes O e P, Sgan 601, Brasília, Distrito Federal 70.830-010, Brazil
[11]Laboratório Central de Saúde Pública do Estado de Mato Grosso (LACEN-MT), Rua Santiago, 70, Cuiabá, Mato Grosso 78.060-628, Brazil
[12]Laboratório Central de Saúde Pública do Estado do Pará (LACEN-PA), Rodovia Augusto Montenegro, 524, Belém, Pará 66823-010, Brazil
[13]Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, Rua Doutor Ovídio Pires de Campos, 225, São Paulo, São Paulo 05403-010, Brazil
[14]Instituto de Biociências, Universidade de São Paulo, Rua do Matão, 321, São Paulo, São Paulo 05508-090, Brazil
[15]Department of Science and Technology for Humans and the Environment, Università Campus Bio-Medico di Roma, Via Alvaro del Portillo, 21, Rome 00128 Rome, Italy
[16]Instituto Rene Rachou, Fundação Oswaldo Cruz, Avenida Augusto de Lima, 1715, Belo Horizonte, Minas Gerais 30190-002, Brazil
[17]Climate Amplified Diseases and Epidemics (CLIMADE), CERI, Tygerberg Medical Campus, Cape Town, South Africa & Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

*Corresponding author. Maria Carolina Elias, Laboratory of Cell Cycle, Instituto Butantan, Av. Vital Brasil, 1500, São Paulo 05503-900, Brazil. E-mail: carolina.eliassabbaga@butantan.gov.br; Sandra Coccuzzo Sampaio, Laboratory of Cell Cycle, Instituto Butantan, Av. Vital Brasil, 1500, São Paulo 05503-900, Brazil. E-mail: sandra.coccuzzo@butantan.gov.brLaboratory; Svetoslav Nanev Slavov, Laboratory of Cell CycleInstituto Butantan, Av. Vital Brasil, 1500, São Paulo 05503-900, Brazil. E-mail: svetoslav.slavov@fundacaobutantan.org.br

Abstract

Influenza A and B viruses represent significant global health threats, contributing substantially to morbidity and mortality rates. However, a comprehensive understanding of the molecular epidemiology of these viruses in Brazil, a continental-size country and a crucial hub for the entry, circulation, and dissemination of influenza viruses within South America, still needs to be improved. This study addresses this gap by consolidating data and samples across all Brazilian macroregions, as part of the Center for Viral Surveillance and Serological Assessment project, together with an extensive number of other Brazilian sequences provided by a public database during the epidemic seasons spanning 2021–23. Phylogenetic analysis of the hemagglutinin segment of influenza A/H1N1pdm09, A/H3N2, and influenza B/Victoria-lineage viruses revealed that in 2021 and in the first semester of 2022, the A/H3N2 2a.3 strain was the predominant circulating strain. Subsequently, the A/H3N2 2b became the prevalent strain until October, when it was substituted by A/H1N1pdm09 5a.2a and 5a.2a.1 lineages. This scenario was maintained during the year of 2023. B/Victoria emerged and circulated at low levels between December 2021 and September 2022 and then became coprevalent with A/H1N1pdm09 5a.2a and 5a.2a.1 lineages. The comparison between the vaccine strain A/Darwin/9/2021 and circulating viruses revealed shared mutations to aspartic acid at residues 186

and 225 across all A/H3N2 lineages from 2021 to 2023, altering the charge in the receptor-binding domain. For A/H1N1pdm09, the 2022 consensus of 5a.2a.1 and the vaccine strain A/Victoria/2570/2019 showed 14 amino acid substitutions. Key residues H180, D187, K219, R223, E224, and T133 are involved in hydrogen interactions with sialic acids, while N130, K142, and D222 may contribute to distance interactions based on docking analyses. Importantly, distinct influenza A lineage frequency patterns were observed across Brazil's macroregions, underscoring the regional variations in virus circulation. This study characterizes influenza A and B viruses circulating in Brazil, providing insights into their circulation patterns and dynamics across Brazilian macroregions. These findings hold significant implications for public health interventions, informing strategies to mitigate transmission risks, optimize vaccination efforts, and enhance outbreak control measures.

Keywords: influenza A; influenza B; genomic surveillance; molecular docking; phylogeny; protein structure prediction; vaccine strains

Introduction

Influenza A and B viruses cause annual worldwide epidemics, significantly impacting public health due to their high morbidity and mortality rates (Keech and Beardsworth 2008, Nair et al. 2011). These infections predominantly manifest as respiratory illnesses, exhibiting varying severity and serious complications, particularly among vulnerable demographics such as children and the elderly (Olson et al. 2007, McCullers and Hayden 2012, Glezen et al. 2013). Annually, influenza is responsible for >200 000 deaths globally (Thompson et al. 2009, Paget et al. 2019, Troeger et al. 2019).

In Brazil, as well as globally, the primary causative agents of seasonal influenza are influenza A subtypes, particularly H1N1 and H3N2 (Petrova and Russell 2018, BRASIL 2023a). However, since 2001, there has been a cocirculation of influenza B lineages, Victoria and Yamagata, with the Victoria lineage emerging as the predominant strain in recent seasons (Koutsakos et al. 2016, Costa et al. 2022).

Previous studies have indicated significant variation in the evolutionary patterns of influenza A and B viruses based on viral strain (Irving et al. 2012, Sharma et al. 2020). While influenza A/H3N2 subtypes exhibit relatively rapid evolution with frequent replacements occurring every 2–5 years, A/H1N1pdm09 and influenza B viruses evolve more slowly (Smith et al. 2004). Despite this slower evolution, the cocirculation of multiple lineages allows for the emergence of new antigenic variants approximately every 3–8 years (Chen and Holmes 2008, Bedford et al. 2014).

Furthermore, these studies underscore the critical role of genomic surveillance of influenza for understanding its epidemiology, incidence, and phylogenetic relationships, which may impact disease severity (Forleo-Neto et al. 2003, Cantarino and Merchan-Hamann 2016). Hemagglutinin (HA), the primary antigenic target of influenza vaccines, plays a crucial role in the pathogenesis of the infection and is subject to antigenic changes that can affect vaccine efficacy (Petrova and Russell 2018, Yamayoshi and Kawaoka 2019, Wu and Wilson 2020).

Despite Brazil's established protocol for influenza sample collection and vaccine strain selection (BRASIL 2023a), there are still gaps in the comprehensive genetic and phylogenetic characterization of influenza A and B viruses in the country. Therefore, this study aims to address this gap by assessing and characterizing influenza A and B viruses collected in the five Brazilian macroregions between 2021 and 2023 through phylogenetic analysis of the HA gene.

Material and methods

Center for Viral Surveillance and Serological Assessment dataset

Clinical sample collection

The Center for Viral Surveillance and Serological Assessment (CeVIVAS) project, a collaborative initiative involving the Instituto Butantan, Hemocentro Ribeirão Preto, several Central Public Health Laboratories (LACEN) across Brazil, and the municipalities of São Paulo and São Bernardo do Campo, aims to enhance viral surveillance and serological assessment. To ensure representative sampling of influenza, CeVIVAS included LACENs from all five Brazilian macroregions: Pará (North), Alagoas and Bahia (Northeast), Federal District (Brasília), Mato Grosso (Midwest), Ribeirão Preto, São Paulo, and São Bernardo do Campo cities (Southeast), and Paraná (South).

Only samples that tested positive for influenza viruses with Ct values of <30 accompanied by available epidemiological metadata, such as the date and sample collection location, were selected and sequenced. This stringent criterion ensures the inclusion of informative samples for a more comprehensive assessment of the prevalence and genetic diversity of circulating influenza A and B viruses in Brazil.

This comprehensive collection effort yielded 1277 newly generated Brazilian HA sequences, distributed as follows: 313 samples of influenza A/H1N1pdm09, 56 from Alagoas, 4 from Bahia, 50 from the Federal District, 23 from Mato Grosso, and 180 from São Paulo; 713 samples of influenza A/H3N2, with contributions from Alagoas (126), Bahia (235), Pará (22), the Federal District (86), Mato Grosso (69), and São Paulo (175); and 252 samples of influenza B/Victoria lineage, 130 from Alagoas and 122 from São Paulo. It is noteworthy that all sequences, but one, exhibit a minimal 83% coverage and complete HA coding DNA sequence (CDS). For further details, refer to Supplementary Fig. S1 and Supplementary Table S1. All sequences generated in this study were deposited in the Global Initiative on Sharing All Influenza (GISAID) database.

The sampling plan was meticulously designed to ensure precision and representativeness of the collected data. We utilized a margin of error (alpha error) of 5% and a confidence interval (CI) of 95%, allowing for precise estimation of population parameters. Additionally, we established a sampling power of at least 80%, ensuring the study's ability to detect true differences or effects if present in the population. The final calculated sample size was increased by 20% to account for possible losses.

Furthermore, acknowledging the influence of seasonality on influenza prevalence, we adjusted the sample to seasonal variation. During high prevalence months (October–May), we reduced the sample by 0.25%, and in the remaining months, we increased it by 0.5% (BRASIL 2021b). This approach ensured adequate sample size in each phase of the study, maintaining the representativeness of the collected data. The sample was calculated at the state level to guarantee precision and representativeness in each macroregion. Furthermore, the sample selection involved municipal diversity within each state.

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Ethics Committee of the Faculty of Medicine of Ribeirão Preto, University of São Paulo (Certificate of Presentation of Ethical Appreciation: 68586623.0.0000.0068).

Influenza A and B whole-genome sequencing

The positive clinical samples for influenza A and B were extracted following the methodology described in Supplementary material. Influenza A genomic sequences were obtained using universal primers for influenza A (Opti1-F1, Opti1-F2, and Opti1-R1) as previously described by Mena et al. (2016) with minor modifications. For specific details on the reaction conditions, refer to Supplementary material. Influenza B genomes were obtained using the same one-step amplification strategy however using a set of influenza B universal primers described by Zhou et al. (2014). Pooled libraries were sequenced using 2 × 150 bp pair-end flow cell kits for NextSeq2000 or 2 × 150 bp pair-end flow cell kits for MiSeq (Illumina). For more information about genomic library preparation and influenza next-generation sequencing, see Supplementary material.

Genome assembly pipeline

The raw reads were submitted to an assembly pipeline consisting of three main steps, as depicted in Supplementary Fig. S2. Initially, Trimmomatic was employed to exclude low-quality reads, adapters, and primer sequences (Bolger et al. 2014).

Subsequently, the filtered reads underwent the first main step of the pipeline, where VAPOR (Southgate et al. 2020) and an influenza genome database were utilized. The influenza genome database comprises sequences sourced from GISAID and GenBank, provided by INSaFLU (Borges et al. 2018). The objective was to identify the best reference sequence for each segment of the influenza virus. To achieve this, the segment sequence with the highest score was extracted from the database using seqtk (Shen et al. 2016, available at https://github.com/lh3/seqtk). This selected reference segment was then employed in the second main step, which involved individual assembly of each genomic viral segment.

The filtered reads were mapped to the selected reference segment using Bowtie2, with the –very-sensitive parameter for segment-specific reads selection (Langmead and Salzberg 2012). Properly paired reads were extracted using SAMtools and BEDtools (Quinlan and Hall 2010, Li and Birol 2018). The segment-specific reads were subsequently submitted to assembly using SPAdes (Bankevich et al. 2012). The refinement process entailed mapping the scaffolds to the selected segment reference using Minimap2 with default parameters (Li and Birol 2018). A pre-consensus was then generated using SAMtools and iVar (Grubaugh et al. 2019). An in-house Python script was applied to substitute degenerated bases inserted by iVar for Ns, as the assembled scaffolds may contain potential Single Nucleotide Polymorphisms.

To obtain the final polished consensus segment, the filtered reads were mapped against the pre-consensus using Bowtie2. The iVar tool was employed with the parameters -m 10 and -q 20, requiring a minimum depth of 10 reads and a frequency of 25%. All the assembled segments from the sample were combined to generate a final genome FASTA file. In cases where segment 4, which codes for the HA gene, was successfully assembled, it was subjected to clade attribution using Nextclade (Hadfield et al. 2018). The clade attribution was performed by comparing the assembled segment against the following references: H1, A/Wisconsin/588/2019 (MW626062); H3, A/Darwin/6/2021 (EPI1857216); Victoria, B/Brisbane/60/2008 (KX058884); and Yamagata, B/Wisconsin/01/2010 (JN993010).

GISAID dataset

To expand our dataset, we obtained additional HA gene sequences from the GISAID EpiFlu Database (Shu and McCauley 2017, available at https://www.gisaid.org/, accessed September 2023). This database includes sequences from Brazil and worldwide, from 2021 to 2023. We concentrated our analysis on the influenza B/Victoria lineage and the influenza A virus subtypes A/H3N2 and A/H1N1pdm09, given the absence of the influenza B/Yamagata lineage detections since 2019 (Costa et al. 2022). The dataset from Brazil, encompassing all macroregions, included 604 HA sequences for influenza A/H1N1pdm09, 1637 for A/H3N2, and 527 for B/Victoria. To perform a comprehensive phylogenetic analysis and minimize bias, we sample non-Brazilian sequences across regions, including Argentina, Australia, China, South Africa, the USA, and the UK (Supplementary Table S3). These locations were carefully chosen to encompass the world's major geographic subdivisions and to ensure the availability of deposited sequences. Argentina, Australia, and South Africa were selected to represent countries in the Southern Hemisphere (SH), spanning the continents of South America, Oceania, and Africa, respectively. Conversely, China, the USA, and the UK were chosen to represent countries in the Northern Hemisphere covering Asia, North America, and Europe, respectively (Chan et al. 2010). We selected only high-quality sequences, with at least 85% coverage and complete HA CDS gene. For detailed information, see Supplementary Table S1.

Also, we selected HA sequences from World Health Organization (WHO)-recommended trivalent egg-based vaccine strains for the 2021–24 epidemic seasons in the SH (WHO 2020b, 2021b, 2022b, 2023b). Specifically, for the influenza A/H1N1pdm09 dataset, the strains A/Victoria/2570/2019, A/Sydney/5/2021, and A/Victoria/4897/2022 were included. For the influenza A/H3N2 dataset, we added A/Hong Kong/2671/2019, A/Darwin/9/2021, and A/Thailand/8/2022. In the dataset for the influenza B/Victoria lineage, we included B/Washington/02/2019 and B/Austria/1359417/2021. Detailed information about these vaccine strains can be found in Supplementary Table S4.

For our final dataset, we merged CeVIVAS, GISAID, and vaccine datasets, resulting in 3966, 8609, and 3996 sequences from A/H1N1pdm09, H3N2, and B/Victoria-lineage viruses, respectively. For more information, see Supplementary Table S1.

Genomic surveillance of influenza in Brazil

The number of available HA sequences in Brazil, categorized by viral subtype within the timeframe of this study, was retrieved from the public EpiFlu-GISAID database (available at https://www.gisaid.org/, accessed September 2023), along with those generated during this study. Epidemiological data on laboratory-confirmed positive cases by viral subtype were obtained from the epidemiological bulletins published by the Brazilian Ministry of Health (BRASIL 2022, 2023a, 2023b). The ratio between the number of available HA sequences and confirmed positive cases was used to evaluate three key aspects: genomic coverage, genomic surveillance capacity, and the potential for underestimation.

Phylogenetic analysis

The phylogenetic trees of the A/H1N1pdm09, A/H3N2, and influenza B/Victoria were inferred using the Nextstrain pipeline (Hadfield et al. 2018, available at https://github.com/nextstrain/seasonal-flu). We utilized the Nextstrain Augur tool to

perform subsampling using the parameters –group-by-country and –sequences-per-group. The –group-by-country parameter organizes the sequences by their country of origin, while the –sequences-per-group parameter specifies the maximum number of sequences selected from each country. The values were set based on the maximum number sequences available from Brazil. Clades and subclades are assigned using the collection of signature mutations provided by Nextstrain for each lineage of A/H1N1pdm09, A/H3N2, and B/Victoria. The maximum likelihood method using IQTree (Nguyen et al. 2015) was applied under the general time-reversible model. To statistically support the phylogenetic trees, we applied ultrafast Bootstrap (UFBoot) using 1000 replicates (Minh et al. 2013). UFBoot confidence values >70% were considered as the cut-off for clustering. In the Nextstrain pipeline, during phylogenetic analysis and refining steps, samples that significantly deviate from the molecular clock model are automatically removed, resulting in phylogenetic trees comprising 3517, 8420, and 3194 sequences from A/H1N1pdm09, A/H3N2, and B/Victoria-lineage viruses, respectively. All trees were edited in R using the ggtree package (Yu et al. 2017).

Genetic analysis of Brazilian influenza lineages

Comparison of vaccine strains and Brazilian consensus

To identify the presence of amino acid substitution in the HA segment in Brazilian sequences, we first categorized them based on their subclade/lineage and year of circulation. Next, only complete sequences with 100% coverage were selected, and a consensus sequence was generated by selecting the bases that appeared in at least 50% of the circulating sequences for each subclade/year group. We then compared the consensus sequences obtained from the Brazilian sequences with the vaccine strains recommended by the WHO for the SH in the seasons spanning from 2021 to 2024. These comparisons were performed using amino acid sequences. The HA epitope regions were defined according to the Nextstrain pipeline, which is a widely used tool for the analysis of viral genetic sequences (Aksamentov et al. 2021). This approach aimed to identify any potential amino acid substitutions in the HA segment of Brazilian sequences compared to the vaccine strains, providing insights into the antigenic variability and potential implications for vaccine efficacy.

Selective pressure analysis

To assess the strength of natural selection acting on the HA gene, we employed the Fast, Unconstrained Bayesian AppRoximation (FUBAR) method on the dataset from influenza A and B viruses circulating in Brazil from 2021 to 2023. FUBAR employs a Bayesian approach to estimate nonsynonymous (dN) and synonymous (dS) substitution rates on a per-site basis for a given coding alignment and its corresponding phylogeny. The FUBAR method assumes that the selection pressure at each site remains constant throughout the entire phylogeny (Murrell et al. 2013). Only complete CDS sequences corresponding to the HA gene were retained, and identical sequences were removed. The phylogenetic tree for the three viruses was generated using IQTree (Nguyen et al. 2015) under the optimal evolutionary model (TVM + F + G4) determined by JModelTest, using UFBoot with 10 000 replicates (Minh et al. 2013, Kalyaanamoorthy et al. 2017). Posterior probabilities range from 0 to 1, with values >0.9 indicating strong evidence of positive selection. FUBAR is available through the HYPhy package (Kosakovsky Pond et al. 2005). Default parameters were utilized for the analysis.

Protein structural prediction analysis

The protein structure prediction was performed using Modeller software version 10.4 (Šali and Blundell 1993). To predict the structure of A/H1N1pmd09 HA, we used the PDB6UYN, PDB6HJQ, and PDB5C0S structures as templates. For A/H3N2 HA, we used the PDB4O58 structure as a template. All templates were selected prioritizing identity (>80%), coverage (>80%), resolution (<3.0 Å), and their interaction with antibodies. After predicting the structures, we performed protonation using PDB2PQR v3.0 (Dolinsky et al. 2004). This step adjusts the ionization states of amino acid residues based on the input data, such as the pH value. We used the Propka method to predict the titration state of amino acid residues at pH 7.4 (Olsson et al. 2011, Søndergaard et al. 2011). To perform all electrostatic analyses, we used the APBS v5.0 software (Jurrus et al. 2018). For molecular docking assays, we used the Autodock Vina program (Trott and Olson 2010). We prepared the vaccine strain egg-propagated A/Victoria/2570/2019 and the consensus of subclade 5a.2a.1 from 2022, as well as the sialic acid (alpha-2,3 and alpha-2,6) at pH 7.4. The Gasteiger partial charge assignment was performed using the MGLTools 1.5.6 program (available at https://ccsb.scripps.edu/mgltools). The grid center for molecular docking was set at the RDB domain, with the coordinates $x = -17.345$, $y = -75.700$, and $z = 48.168$. The dimensions of the grid were set as $X = 18$ Å, $Y = 22$ Å, $Z = 18$ Å. This grid represents the region where the docking simulations will be performed to predict the binding affinity between the HA protein (vaccine strain and the consensus from the predominant Brazilian strains) and sialic acids.

Epidemiological profile of influenza viruses and statistical analysis

To assess the distribution and replacement dynamics of influenza A and B virus clades in Brazil before, during, and after the coronavirus disease 2019 (COVID-19) pandemic, we further analyzed HA gene sequences from January 2019 to December 2020. These sequences were retrieved from the public EpiFlu GISAID database (https://www.gisaid.org/, accessed October 2024).

Given that clade diversity across Brazil's regions was only observed among influenza A viruses, we applied the chi-squared test to assess the association between two categorical variables: viral subclades and macroregions (Rana and Singhal 2015). To conduct this analysis, we assumed that the distribution of viral clades across different regions of Brazil is homogeneous. Therefore, the expected occurrence of each viral subclade in each region was estimated based on its overall prevalence across all macroregions. Comparisons were made for distinct circulation periods of influenza A subtypes: H3N2 (November 2021 to November 2022) and H1N1pdm09 (November 2022 to May 2023), considering the entire time frame in the analysis. The expected occurrences of each subclade were multiplied by the total observed values for each macroregion, and the result was divided by the total number of observations within that macroregion. We utilized chi-squared tests with a significance level of .05 to compare the number of viral subclades across Brazil's macroregions. Since the *P*-value was below the threshold, we rejected the null hypothesis. A *post hoc* adherence test was then applied to pinpoint the specific regions where statistically significant differences occurred, with any differences exceeding the critical value considered significant. All statistical analyses were conducted using R.

The sampling power of the combined dataset (GISAID and CeVIVAS) was calculated as described in the section Clinical Sample Collection.

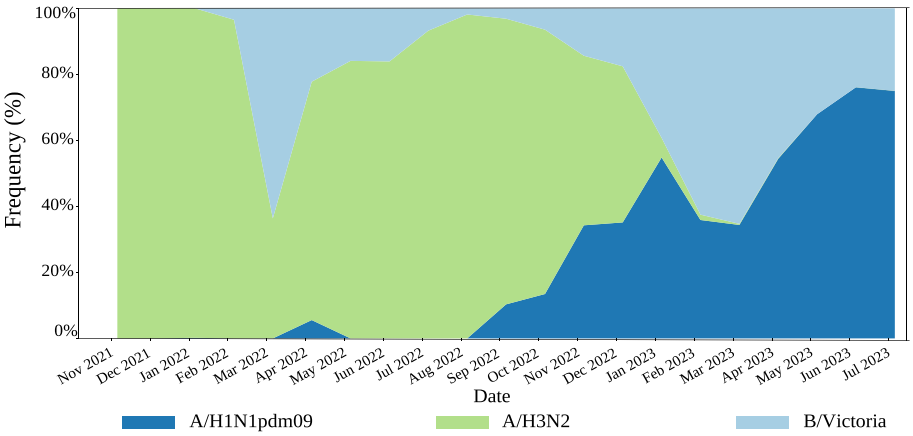

Figure 1. Frequency of influenza viruses circulating in Brazil between 2021 and 2023 according to the performed phylogenetic analysis.

Results

To better understand the panorama of the molecular epidemiology of influenza circulating in Brazil between 2021 and 2023, we generated genome sequences of these viruses from different Brazilian states. We analyzed them together with Brazilian sequences available in the public database. Through analysis of the public data, we could note that the viral sequence data exhibited a steady rise over the last 3 years in Brazil. Still, the ratio of available HA gene sequences concerning positive cases did not exceed 12% (Supplementary Fig. S3).

Our results revealed the circulation of A/H1N1pdm09, A/H3N2, and B/Victoria in Brazil between 2021 and 2023. It was possible to detect the predominance of the A/H3N2 virus in 2021/2022 and the cocirculation of the three viruses at the end of 2022. However, in 2023, our analyses indicated a sharp increase in the circulation of both B/Victoria and A/H1N1pdm09 viruses in the country (Fig. 1).

Additionally, sequences collected from Brazil prior to and during the COVID-19 pandemic exhibited greater diversity compared to the postpandemic period (Supplementary Fig. S4). Following the pandemic, a near-complete replacement of subclades was observed in both influenza A and B clades. Only one subclade each of A/H1N1pdm09 (6B.1A.5a.1) and A/H3N2 (3C.2a1b.2a.2a.3) remained in circulation after the onset of COVID-19 (Supplementary Fig. S4).

Phylogenetic analyses

A/H1N1pdm09—Predominance of viruses belonging to the 5a.2a.1 subclade

Phylogenetically, all 896 Brazilian sequences were grouped in three clade/subclades: almost in 5a.2a.1 ($n = 626$, 70%) and, to a lesser extent, in subclade 5a.2a ($n = 252$, 28%) and clade 5a.1 ($n = 18$, 2%) (Fig. 2a) that were circulating in the period when this study was concluded (first semester of 2023). In the 5a.2a.1 subclade, two main clusters were recovered. The first cluster (I-blue, $n = 37$, 4% of the total) was characterized by T216A amino acid substitution and consisted of sequences primarily from the Southeast region ($n = 32$). These sequences were also grouped as a sister group of the vaccine strain A/Victoria/4897/2022—designated for administration in 2024 in the National Influenza Vaccination Campaign (BRASIL 2024). The second cluster (II-blue, $n = 592$, 65.8% of the total) encompasses most Brazilian sequences from all regions of the country. From that specific cluster, 115 sequences exhibited the T270A amino acid substitution, with a notable prevalence in the North ($n = 61$) and Northeast ($n = 34$) regions (Fig. 2b).

Within subclade 5a.2a, four clusters were recovered: the first cluster (I-pink; $n = 43$, 4.8% of the total) mainly consisted of sequences from the Northern macroregion ($n = 35$) of the country displaying the S83P amino acid substitution (Fig. 2c). Also, this cluster was recovered as a sister group of the vaccine strain A/Sydney/5/2021—administered in 2023 by the National Influenza Vaccination Campaign (BRASIL 2023a). The second cluster (II-pink; $n = 19$, 2.1% of the total) grouped sequences from the Southern ($n = 17$) region and was characterized by A48P amino acid substitution. The third cluster (III-pink; $n = 13$, 1.4% of the total) predominantly gathered sequences from Brazil's Northeast ($n = 9$) region. Lastly, the fourth cluster (IV-pink, $n = 170$, 19% of the total) grouped sequences from all five macroregions of the country.

Within clade 5a.1, all Brazilian sequences ($n = 18$, 2% of the total; Fig. 2d) were grouped into a single clade, characterized by the shared I149V amino acid substitution. This clade predominantly consisted of sequences from the Southern region ($n = 15$) and fewer representatives from the Midwest ($n = 1$) and Southeast regions ($n = 1$).

Selection pressure analysis identified four sites in the HA gene under positive selection within the A/H1N1pdm09 subclades: 137, 156, 163, and 164. When comparing the HA consensus sequences of viruses circulating in Brazil between 2021 and 2022 with the egg-propagated vaccine strain A/Victoria/2570/2019—vaccine administered in the 2021/2022 seasons (BRASIL 2021a, 2022) —8–16 amino acid substitutions were revealed (Supplementary Tables S5–S6). Specifically, when compared with the predominant Brazilian subclade (5a.2a.1) in 2022, 14 amino acid substitutions were detected, 11 of which occurred in epitope regions (Supplementary Table S5). Notably, one of these sites (137) was found to be under positive selection. Additionally, three substitutions (R223Q, D260E, and T277A) are particularly significant due to their positions in the receptor-binding domain (RBD) and the fusion domain, as predicted by protein structure modeling (Fig. 3). In contrast, a general reduction in the number of amino acid substitutions (4–10) was observed when comparing the predominant subclades collected in Brazil in 2023 with the egg-propagated vaccine strain A/Sydney/5/2021, except for the 5a.1 subclade (Supplementary Table S7). Four common amino acid substitutions (N94D, A216T, R223Q,

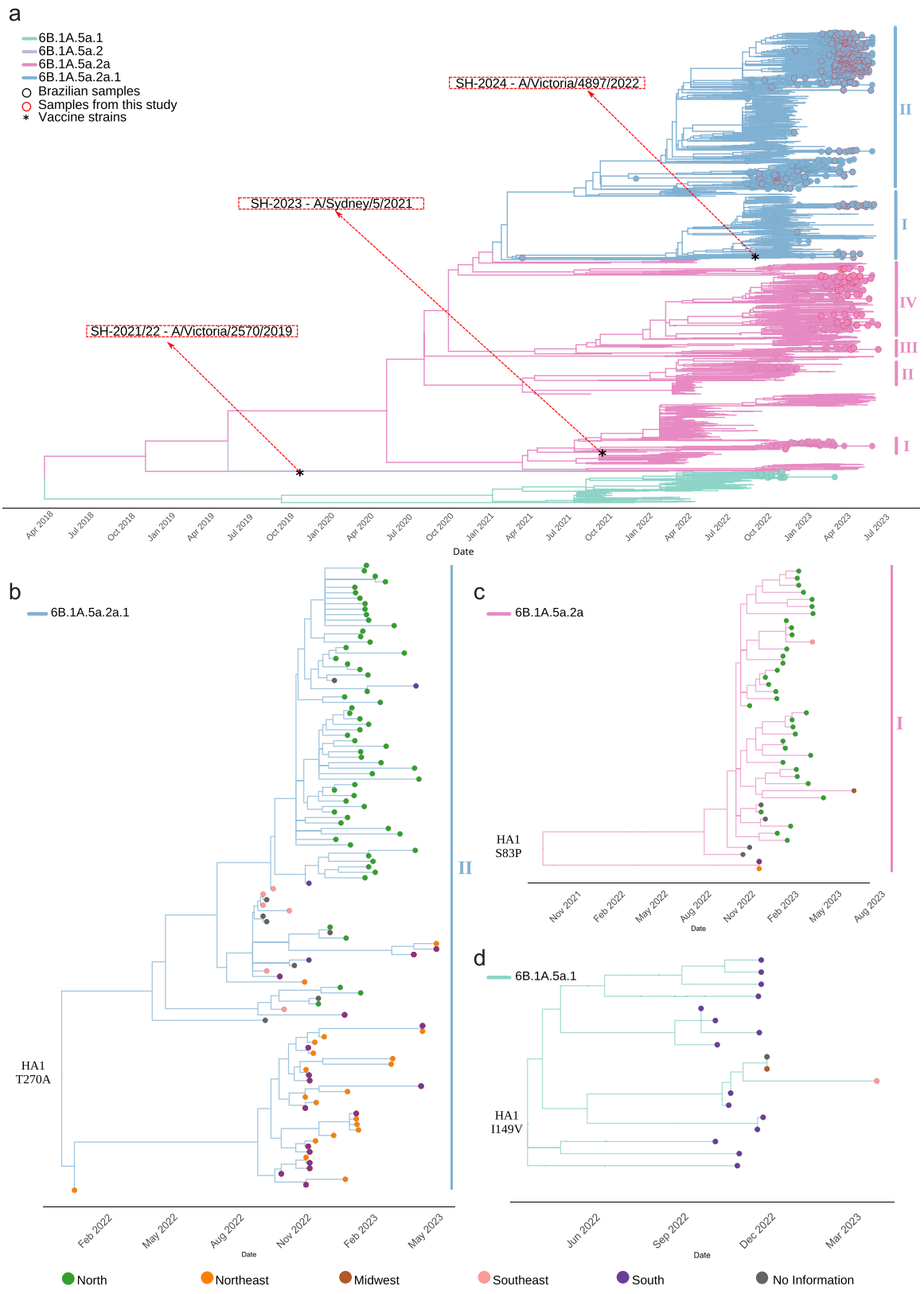

Figure 2. (a) Worldwide time-aware phylogeny of the genomic HA segment of the A/H1N1pdm09 viruses collected between January 2021 and July 2023 constructed by Nextstrain. In Brazil, subclades 5a.2a.1 (clusters I and II) and 5a.2a (clusters I–IV) are notably predominant. The SH recommended vaccine strains from 2021 to 2024, which are indicated by asterisks and red arrows. The black circles identified Brazilian sequences; the red circles identified sequences generated in this study. (b) Zoom-out of 5a.2a.1 II cluster of Brazilian sequences mostly from the North and Northeast regions with HA1 T270A amino acid substitutions. (c) Zoom-out of 5a.2a I cluster consisted of sequences mainly from the North region with HA1 S83P amino acid substitution. (d) Zoom-out of 5a.1 grouping sequences mainly from the South region, characterized by HA1 I149V amino acid substitution.

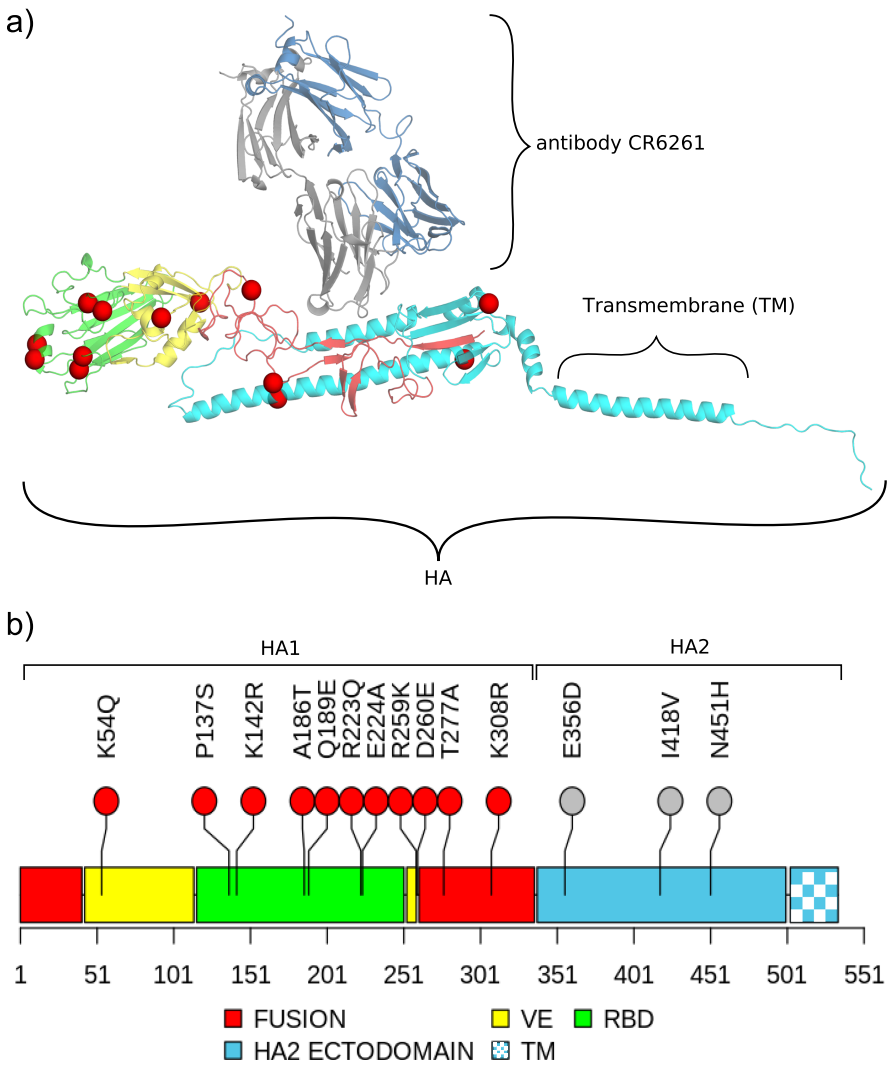

Figure 3. Influenza virus A/H1N1pdm09 HA protein structure and observed mutations between vaccine strain egg-propagated A/Victoria/2570/2019 and Brazilian consensus from subclade 6B.1A5a.2a.1 (5a.2a.1) in 2022. (a) 3D predicted structure of complex antibody, HA, colored by different domains. Fusion domain (red), HA2 ectodomain (cyan), VE domain (yellow), RBD receptor-binding domain (green), and transmembrane (TM) anchor (checkerboard cyan); substitutions in HA epitope regions indicated by red spheres. The antibody is colored by different chains: light chain (gray) and heavy chain (sky blue). (b) 1D schematic view of HA domains, colored by different domains. Color coding as per (a). The main amino acid substitutions are shown in epitope positions (red circle) and nonepitope regions (gray circle).

and I418V) were identified upon comparison of the consensus sequences from subclades 5a.2a.1 and 5a.2a with the strain A/Sydney/5/2021. The R223Q mutation is specifically at the residue responsible for interaction with sialic acid. As presented in Fig. 3a, none of those mutations were close to the region that interacts with some antibodies like CR 6261. However, the 5a.1 subclade exhibited a higher number of substitutions, ranging from 10 to 16, following the vaccine update. Within the epitope region, the K156N substitution was present in comparison to both vaccine strains administered in the country between 2021 and 2023. This site was also identified as being under positive selection in our analysis. For more details, see Supplementary Tables S5–S7.

According to the virtual docking assay, the vaccine strain A/Victoria/2570/2019 RBD residues as H180, D187, K219, R223, E224, and T133 are responsible for hydrogen interactions with alpha-2,3 or alpha-2,6 sialic acids, while N130, K142, and D222 residues could be responsible for distance interactions. When examining the consensus sequence of subclade 5a.2a.1, we noted that only residues D187 and Q223 exhibited hydrogen interactions, while W150 potentially engaged in a CH–pi interaction with alpha-2,3 sialic acid. Regarding alpha-2,6 sialic acid, we observed that residue L191 could partake in a hydrophobic interaction with the substrate, while residues D187 and Q223 persisted in hydrogen interactions. For more detailed information, see Supplementary Fig. S5.

Influenza A/H3N2—high diversification and low circulation during the 2023 season

Phylogenetically, all 2265 Brazilian sequences were clustered mainly within subclade 3C.2a1b.2a.2a.3 (2a.3, $n = 1687$, 74.5%), followed by subclade 3C.2a1b.2a.2b (2b, $n = 456$, 20.1%), and to a lesser extent in subclades 3C.2a1b.2a.2c (2c, $n = 76$, 3.4%), 3C.2a1b.2a.2a (2a.2a, $n = 35$, 1.5%), 3C.2a1b.2a.2a.1 (2a.1, $n = 9$, 0.4%), and 3C.2a1b.2a.2a.1b (2a.1b, $n = 2$, 0.1%) (Fig. 4a).

In our tree, within subclade 2a.3 two clusters were recovered: the first (I-blue, $n = 10$, 0.4%) comprised only six sequences from the Midwest region, and the second cluster (II-blue, $n = 1677$, 74%) encompassed most Brazilian sequences from all regions of

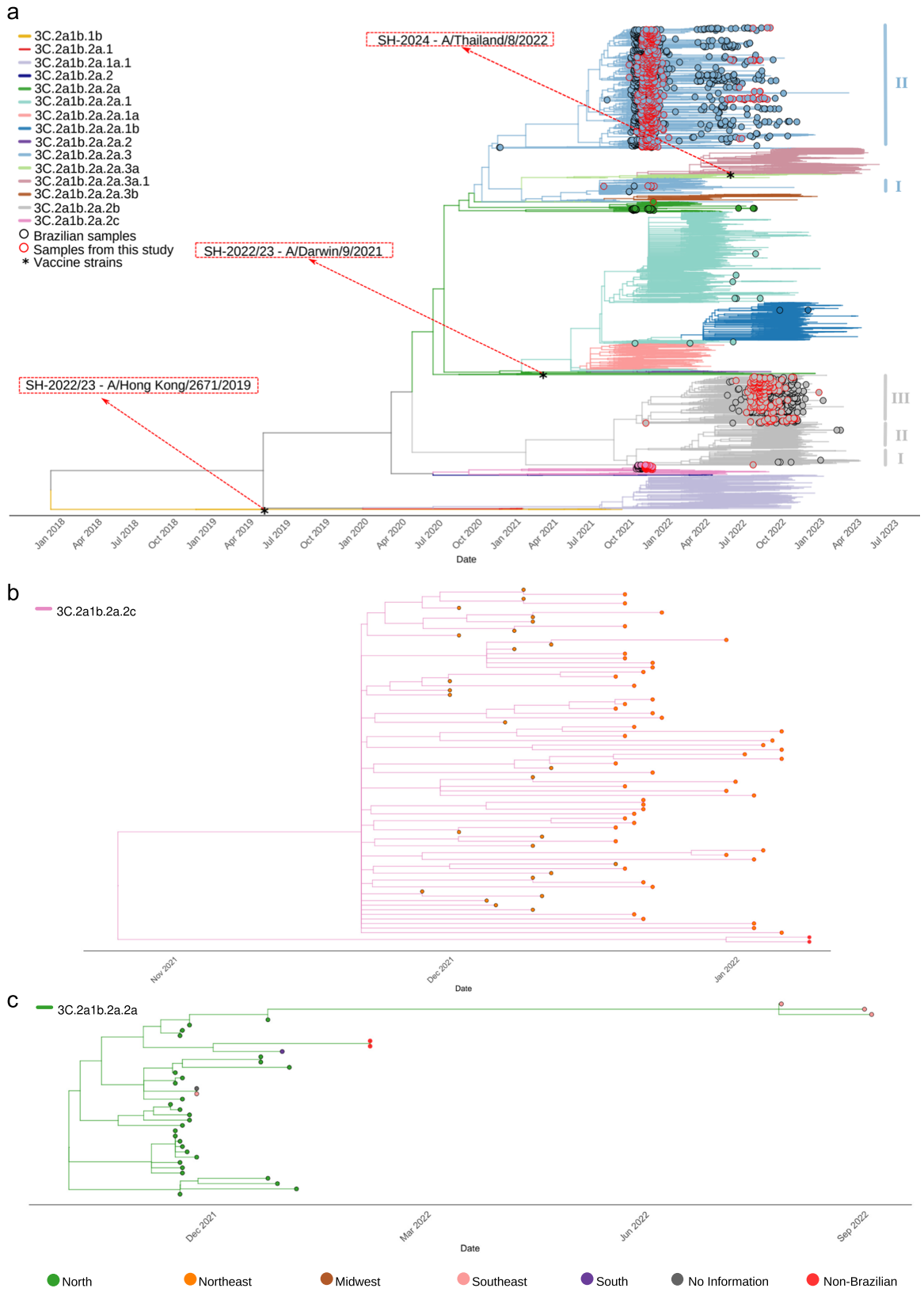

Figure 4. Worldwide time-aware phylogeny of the genomic HA segment of influenza A/H3N2 collected between January 2021 and July 2023 constructed by Nextstrain. (a) In Brazil, predominance and cocirculation of the 2a.3 (clusters I and II blue) and 3C.2a1b.2a.2b (clusters I–III gray) subclades. (b) Zoom-out of subclade 2c, demonstrating exclusively Northeast sequences. (c) Zoom-out of subclade 2a.2a, which was comprised mostly of sequences from the North region. The black circles identified Brazilian sequences; the red circles identified sequences generated in this study. The SH recommended vaccine strains from 2021 to 2024, which are indicated by asterisks and red arrows.

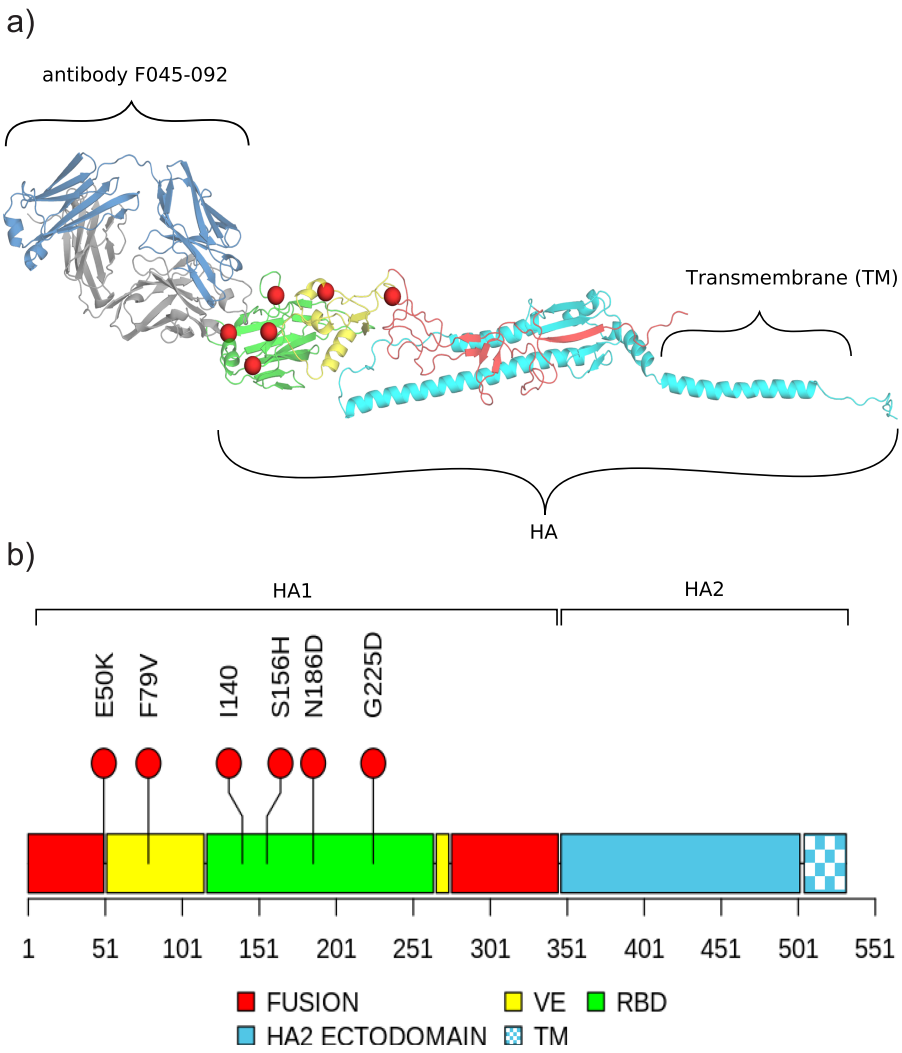

Figure 5. Influenza virus A/H3N2 HA protein structure and observed mutations between vaccine strain egg-propagated A/Darwin/9/2021 and Brazilian consensus from subclade 3C.2a1b.2a.2b (2b) in 2022. (a) 3D predicted structure of complex antibody HA, colored by different domains. Fusion domain (red), HA2 ectodomain (cyan), VE domain (yellow), RBD receptor-binding domain (green), and transmembrane (TM) anchor (checkerboard cyan); substitutions in HA epitope regions indicated by red spheres. The antibody is colored by different chains: light chain (gray), and heavy chain (sky blue). (b) 1D schematic view of HA domains, colored by different domains. Color coding as per (a). The main amino acid substitutions are shown in epitope positions (red circle).

the country. In subclade 2b, three clusters were recovered: the first cluster (I-gray, $n = 4$, 0.2%) consisted of sequences originating from the South ($n = 3$) and North ($n = 1$) regions sharing the I242M amino acid substitution. In the second cluster (II-gray, $n = 2$, 0.8%), the sequences from the South ($n = 2$) region presented S262N, T135A amino acid substitution. Within the third cluster (III-gray, $n = 449$, 19.8%), most sequences originated from all regions of the country. The 2c subclade consisted exclusively of Northeast sequences ($n = 76$, 3.3%, Fig. 4b). Also, this subclade was recovered as the sister group of the 3a.1, which includes the vaccine strain Thailand/8/2022—designated for administration in 2024 in the National Influenza Vaccination Campaign (BRASIL 2024). In the 2a.2a subclade, most of the sequences were from the North ($n = 35$, 1.5%, Fig. 4c). In the 2a.1 subclade, sequences from all regions except the North were grouped. Within the 2a.1b subclade, only sequences from the Northeast ($n = 2$) were recovered.

Selection pressure analysis identified two sites in the HA gene under positive selection within the A/H3N2 subclades: 53 and 378. The comparison between the HA consensus of viruses circulating during 2021 in Brazil and the vaccine strain egg-propagated A/Hong Kong/2671/2019, which constituted the vaccine administered in the 2021 season (BRASIL 2021a), revealed 18–21 amino acid substitutions (Supplementary Table S8). In contrast, fewer amino acid substitutions (5–8) were observed when comparing viruses collected in 2022 in Brazil with the vaccine strain egg-propagated A/Darwin/9/2021, which was administered from April 2022 to May 2023 during the National Influenza Vaccination Campaign (BRASIL 2022) (Supplementary Table S9). Specifically, when comparing the predominant Brazilian subclade (2a.3) from 2021 to early 2022, the vaccine update resulted in a reduction of amino acid substitutions, decreasing from 21 to 7 (Supplementary Table S10). Mutations at sites 53 and 378 were consistently observed when comparing this lineage with both vaccine strains administered in the country between 2021 and 2023. These sites were identified as being under positive selection in our analysis. Upon comparing A/Darwin/9/2021 and the prevailing Brazilian subclade (2b) during the second semester of 2022, our analysis unveiled

six amino acid substitutions, all localized within epitope regions, which of them stands out the mutations at sites 156, 186, and 225 (Fig. 5). These mutations are located within the sialic acid binding site, and both mutations at sites 186 and 225 were identified across all A/H3N2 virus lineages circulating in Brazil between 2021 and 2023. These mutations resulted in a change of charge from an uncharged polar residue (Asparagine—N or Glycine—G) to a negatively charged residue (Aspartate—D). On the other hand, the mutation at site 156 resulted in a change of charge from an uncharged polar residue (Serine—S) to a positively charged residue (Histidine—H). The latter mutation (S156H) was also found within the viruses from the 2c and 2a.1b subclades. For more details, see Supplementary Tables S8–S10.

Electrostatic analyses indicated that E50K and F79V amino acid substitutions are in the fusion (F′) and vestigial esterase (VE) domains of HA1, respectively (Fig. 5). The E50K substitution changes the electrostatic profile, while F79V could promote a change in the orientation of the side chain of amino acids close to it (Supplementary Fig. S6). The I140K, N186D, and G225D mutations induce a little difference in the electrostatic profile of the RBD domain, making it a little more electro-negative to the RBD domain of the vaccine strain (Supplementary Fig. S7). As presented in Fig. 5a, all those mutations were close to the region that interacts with some antibodies like F045-092.

Influenza B/Victoria-lineage viruses—high circulation in 2023 with regional arrangement

The obtained phylogenetic tree revealed that all 778 Brazilian influenza B viruses were classified as the Victoria lineage from the V1A.3a.2 clade (Fig. 6a) and among those, almost all presented the D197E amino acid substitution ($n = 649$, 83.4%). Within the V1A.3a.2 clade, nine clusters were recovered. The first (I, $n = 42$, 5.4%) was composed mostly of sequences from the North ($n = 40$) region of Brazil, in which T182A, D197E, and T221A were characterized (Fig. 6b). The second cluster (II, $n = 85$, 11%) encompasses sequences mostly from the Northeast, and a few sequences from all regions of the country (Fig. 6c). This cluster was characterized by S208P, E128K, and A154E substitutions. The third (III) cluster encompassed two sequences from the Northeast and shared the G141R substitution. Also, this cluster was recovered as a sister group of the recommended vaccine strain B/Austria/1359417/2021—which made up the vaccine administered during the 2022 and 2024 seasons (BRASIL 2022, 2023a, 2024). The fourth (IV) cluster comprised only one sequence from the Northeast region and displayed D197E, V117I, A154T, K326R, and 128K amino acid substitution. The fifth cluster (V, $n = 104$, 13.4%) grouped sequences from all regions and can be characterized by D129G, E183K, and V87A substitutions. The sixth cluster (VI, $n = 83$, 10.7%) grouped sequences ($n = 83$) mostly from the Northeast ($n = 31$) and North ($n = 30$) regions, which shared the HA1 D197E and Q200P amino acid substitutions (Fig. 6d). The seventh (VII) cluster contains most of the Brazilian sequences ($n = 315$), of which almost all come from the Southeast and Northeast regions, both represented in nearly equal proportions. The eighth cluster (VIII, $n = 10$, 1.3%) were grouped sequences from all Brazilian regions, except the South, and presented the D197E and E183K substitutions. In the ninth cluster (IX, $n = 107$, 13.7%), almost all sequences were from the Northeast ($n = 97$) region, and some were from the Southeast ($n = 8$) and North ($n = 2$) regions. This cluster was characterized only by the D197E amino acid substitution.

The selection pressure analysis identified three sites in the HA gene under positive selection within the B/Victoria viruses: 80, 154, and 410. When comparing the HA consensus sequences of the viruses circulating during 2022 with the vaccine strain egg-propagated B/Austria/1359417/2021, only two amino acid substitutions were detected (D197E and Q200P) (Supplementary Tables S11–S12). In contrast, when examining viruses collected in Brazil in 2023, only one amino acid substitution (D197E) was observed (Supplementary Table S11). Notably, none of the sites under positive selection exhibited substitutions.

Epidemiological profile of influenza A viruses in macroregions of Brazil

Both phylogenetic analysis and the chi-squared test suggested dependence in the distribution of influenza subclades across the Brazilian territory. Given Brazil's diverse climate and demographic characteristics across its five macroregions, we sought to determine whether distinct circulation patterns exist between these regions. First, we assessed whether the data were sufficient for such a comparison. The analysis showed that A/H3N2 sequences, predominating between November 2021 and November 2022, provided a sampling power of 98% (CI 97%), while A/H1N1pdm09 sequences, predominating between November 2022 and May 2023, yielded a sampling power of 87% (CI 90%). Subsequently, we conducted subsampling of the final dataset, focusing on peak periods for both A/H1N1pdm09 and A/H3N2 subtypes, to investigate the nationwide epidemiological profile of influenza A. Concerning A/H3N2 subclades (Fig. 7), the chi-squared analysis revealed a significant difference in lineage occurrences across regions ($\chi^2 = 513.6$, df = 24, $P < 2.2e\text{-}16$, critical value = ± 1.49). We observed a significantly higher frequency of subclade 2a.3 in the Midwest compared to the expected value (critical value = 3.49), while subclade 2b exhibited higher circulation, particularly in the Southeast (critical value = 6.43). Subclade 2a was detected in three regions, with the North showing significantly higher circulation than expected (critical value = 16.11). Additionally, the 2c subclade was found exclusively in the Northeast (critical value = 14.30). For detailed regional diversity, see Supplementary Fig. S8, and for statistical information, see Supplementary Table S13.

As with A/H3N2, the chi-squared analysis for A/H1N1pdm09 (Fig. 8) also showed a significant difference in lineage occurrences across regions ($\chi^2 = 209.76$, df = 8, $P < 2.2e\text{-}16$, critical value = ± 1.35). According to the results, the 5a.2a.1 subclade predominated in the Northeast, Midwest, and Southeast macroregions, where it circulated at a higher frequency than expected (critical value ≤ 2.20). In the North, the 5a.2a subclade circulated more widely (critical value = 9.96). Although the 5a.1 subclade was infrequent nationally, higher circulation was observed in the South (critical value = 10.38). For detailed regional diversity see Supplementary Fig. S9, and for detailed statistical information, see Supplementary Table S14.

Discussion

The evaluation of lineage diversity among seasonal influenza viruses circulating in Brazil from 2019 to 2023, as shown in this study, reflects the global decline in genetic and antigenic diversity during this period (Dhanasekaran et al. 2022). Typically, seasonal influenza viruses sustain their diversity through continuous reintroduction from tropical regions and opposite hemispheres (Petrova and Russell 2018). However, the COVID-19 pandemic restrictions—such as mask-wearing, travel limitations, and social distancing—significantly reduced global circulation, disrupting this process and resulting in a marked worldwide decline in viral diversity (Dhanasekaran et al. 2022). Following the pandemic, influenza activity in Brazil surged, with positive cases rising from

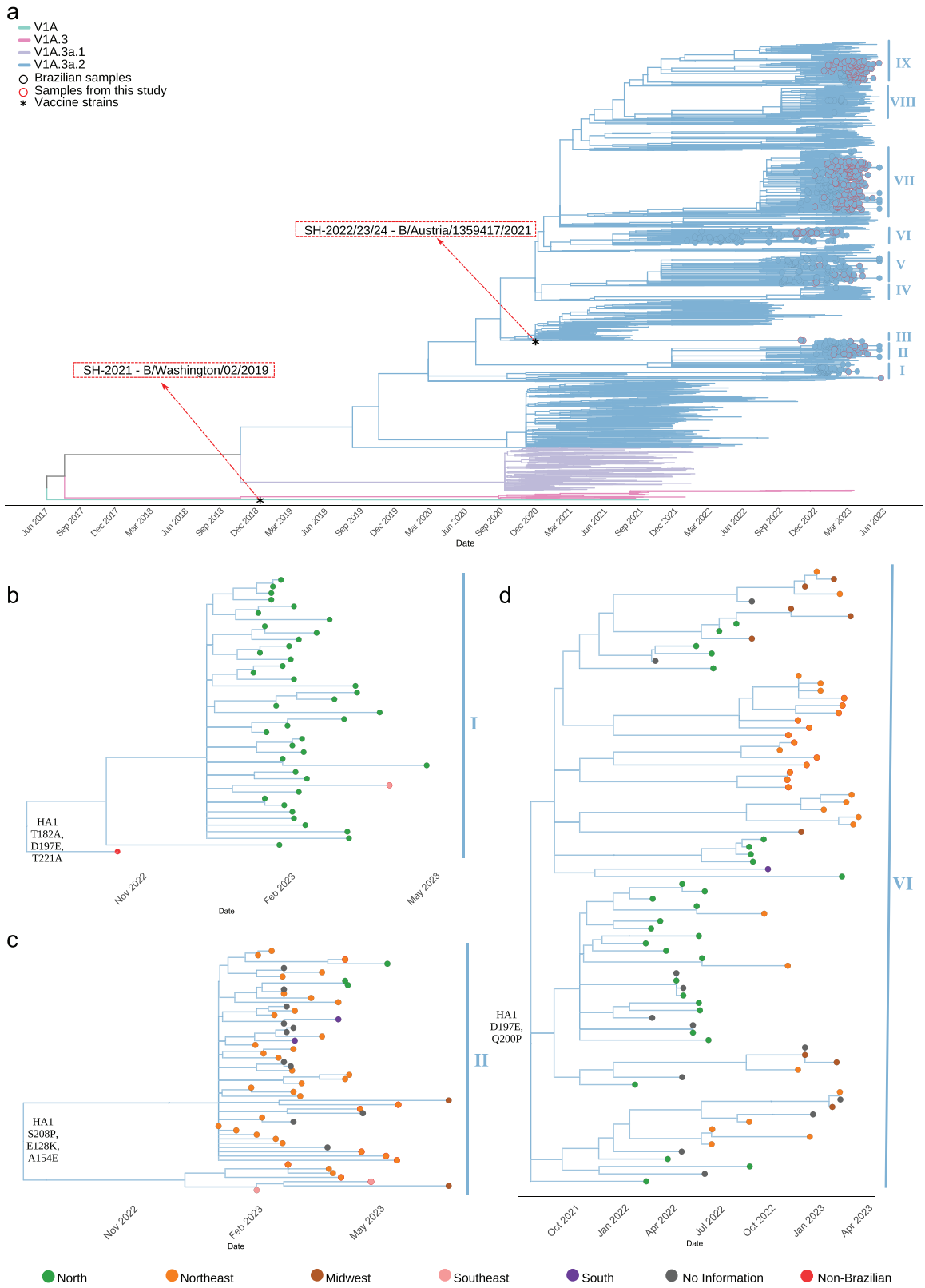

Figure 6. Worldwide time-aware phylogeny of the genomic HA segment of the influenza B/Victoria-lineage viruses collected between January 2021 and July 2023 constructed by Nextstrain. (a) In Brazil, it was exclusively detected the circulation of the V1A.3a2 subclade. (b) Zoom-out of cluster I (characterized by HA1 T182A, D197E, and T221A amino acid substitutions) composed mostly of sequences from the North region. (c) Zoom-out of cluster II comprises mostly sequences from the Northeast, with S208P, E128K, and A154E amino acid substitutions. (d) Zoom-out of cluster VI, grouping sequences mostly from the Northeast and North regions, sharing HA1 D197E and Q200P amino acid substitutions. The black circles identify Brazilian sequences; the red circles identify sequences generated in this study. A red circle identifies the sequences generated in this study. The SH recommended vaccine strains from 2021 to 2024, which are indicated by asterisks and red arrows.

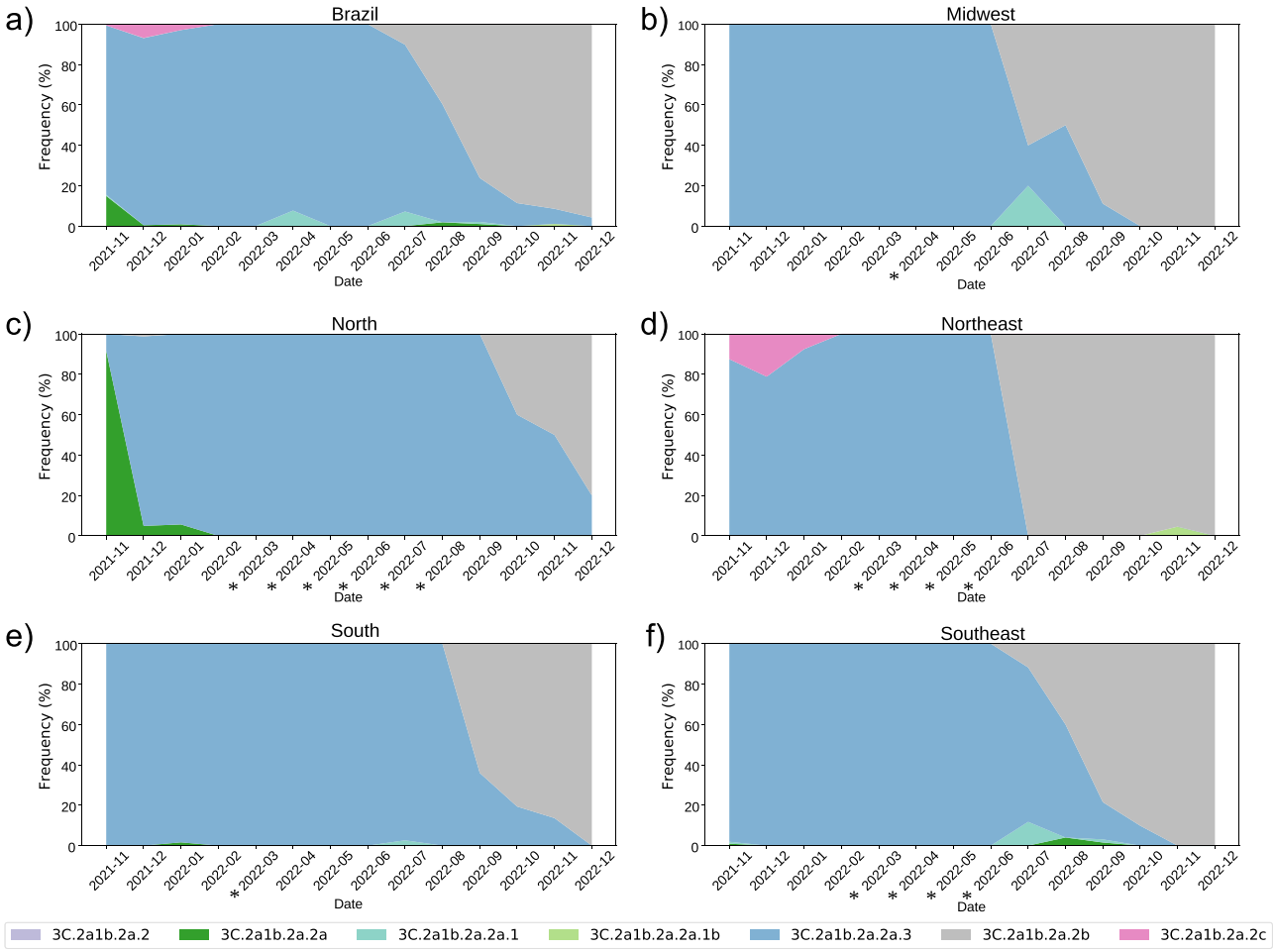

Figure 7. Frequency of A/H3N2 subclades 3C.2a1b.2a.2a.1 (2a.1), 3C.2a1b.2a.2a.3 (2a.3), 3C.2a1b.2a.2a.2b (2b), 3C.2a1b.2a.2 (2), 3C.2a1b.2a.2a (2a), and 3C.2a1b.2a.2c (2c) circulating from November 2021 to November 2022 in Brazil. (a) Frequency of A/H3N2 subclades in Brazil. (b) Frequency of A/H3N2 subclades in Midwest. (c) Frequency of A/H3N2 subclades in North. (d) Frequency of A/H3N2 subclades in Northeast. (e) Frequency of A/H3N2 subclades in South. (f) Frequency of A/H3N2 subclades in Southeast. The asterisk indicates the period of absence from genomic monitoring.

5252 in 2021 to 18 736 in 2023, highlighting a notable resurgence (BRASIL 2022, 2023a, 2023b). Notably, between September 2021 and January 2022, influenza viruses circulated at significantly reduced rates, comprising <3% of cases, compared to the prepandemic rate of 17% (WHO 2022a). The decrease in influenza circulation has been linked to the widespread disruption caused by the COVID-19 pandemic (Dhanasekaran et al. 2022). This global health crisis prompted a concerted effort to enhance molecular surveillance of pathogen genomes, emerging as a critical measure for preparedness against future epidemics (Attwood et al. 2022, Ladner and Sahl 2023). The progressive sequencing efforts spanning from 2021 to 2023 in Brazil reflect this effort and have unveiled the identification of three influenza viruses circulating in Brazilian territory: A/H1N1pdm09, A/H3N2, and B/Victoria lineage. This finding is consistent with data from the Brazilian Epidemiological Surveillance (BRASIL 2021a, 2022, 2023b) and further supports the global cocirculation of influenza A and B viruses.

Our phylogenetic analysis of A/H1N1pdm09 revealed a cocirculation of subclades 5a.2a.1, 5a.2a, and 5a.1. The predominance of 5a.2a.1 viruses across Brazil is in line with observations from other regions, including North America and Europe (WHO 2023a). Notably, the HA1 T216A amino acid substitution, found in some sequences from the Southeast, has been previously identified in many 5a.2a.1 viruses (WHO 2023a). Additionally, the T270A amino acid substitution observed in sequences from the North and Northeast regions appears as a recurrent mutation within this subclade, although its functional significance remains uncertain (Tapia et al. 2020, Zolotarova et al. 2021).

Among the four clusters recovered within 5a.2a, three of them were associated with specific Brazilian macroregions, which may suggest regional characteristics of the circulating influenza A strains. For instance, a cluster composed of sequences obtained from the North Brazil was characterized exclusively by the presence of the S83P substitution that has also been described in sequences obtained from Asia, i.e. Thailand (ECDC 2023b), and the Middle East as well as Oceania (available at https://nextstrain.org/flu/seasonal/h1n1pdm/ha/2y?branchLabel=aa). While this might be an indication of a possible introductory route from these locations, to confirm such a hypothesis more specialized phylogeographic analysis is warranted. This approach is essential for elucidating how distinct regions influence the introduction and dissemination of the virus throughout Brazil.

Phylogenetic analysis revealed a close relationship between A/H1N1pdm09 Brazilian viruses and the vaccine strains. Furthermore, the decrease in amino acid substitutions from 14 to 10 in circulating viruses in 2023, when compared to the vaccine strain A/Sydney/5/2021, suggests a low level of antigenic drift. This reduction may be the result of the selection and updating

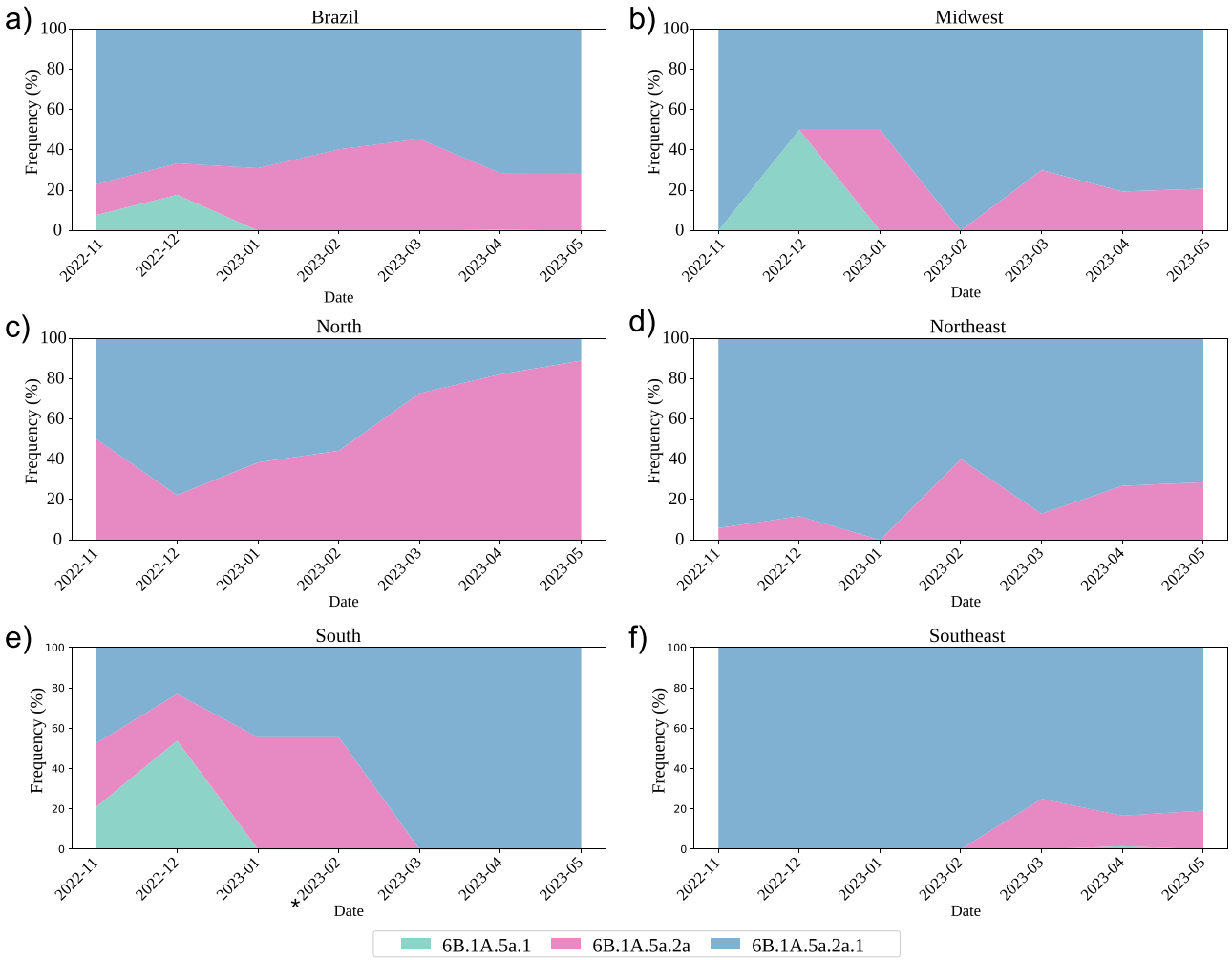

Figure 8. Frequency of A/H1N1pdm09 subclades 6B.1A.5a.1 (5a.1), 6B.1A.5a.2a (5a.2a), and 6B.1A.5a.2a.1 (5a.2a.1) from November 2022 to May 2023 in Brazil. (a) Circulation in Brazil. (b) Frequency of A/H1N1pdm09 subclades in Midwest. (c) Frequency of A/H1N1pdm09 subclades in North. (d) Frequency of A/H1N1pdm09 subclades in Northeast. (e) Frequency of A/H1N1pdm09 subclades in South. (f) Frequency of A/H1N1pdm09 subclades in Southeast. The asterisk indicates the period of absence from genomic monitoring.

of appropriate vaccine strains. Among the amino acid substitutions at antigenic sites, P137S, K142R, D260E, and T277A substitutions are characteristics of viruses belonging to the 5a.2a.1 subclade (WHO 2022b). These substitutions were related to reduced geometric mean titers postvaccination, especially to the egg-propagated vaccine strain A/Sydney/5/2021 (WHO 2023a). The site 137 is under positive selection and may explain the partial immune evasion. The SH vaccine A/Victoria/4897/2022 covered the circulating clades (WHO 2023a).

A previous study highlights the complex evolution of A/H3N2, characterized by multiple clades globally distributed, posing challenges for vaccine composition (Pillai et al. 2023). Our analysis identified seven clades/subclades circulating in Brazil, with the 2a.3 and 2b subclades being more prevalent. Within the 2a.3 subclade, Brazilian sequences grouped into two clusters, while the 2b subclade exhibited three clusters, one of which comprised 99% of sequences from all regions of Brazil. Despite the low circulation of viruses from the remaining 2a.2a, 2c, and 2a.1b subclades, they were restricted to the North and Northeast regions. The diversity of A/H3N2 viruses elucidated through our analysis corroborates the findings outlined in the WHO reports spanning the years 2020–22 (WHO 2020a, 2020b, 2021a, 2021b, 2022b). These findings align with WHO reports and underscore H3's rapid evolutionary rate (Fitch et al. 1997, Bush et al. 1999, Ferguson et al. 2003).

More than 17, and up to a maximum of 21, amino acid substitutions were detected when comparing the vaccine strain A/Hong Kong/2671/2019 (2021 season) with the subclades circulating in Brazil. The predominant 2021 subclade (2a.3) accumulated 21 amino acid substitutions, 17 of which are located within major antigenic sites, including E53N, K135T, and H156S. The H156S and D53N substitutions are characteristic of the 2a.3 subclade (WHO 2023b), with site 53 identified as being under positive selection pressure. While the high number of mismatches alone is insufficient to infer reduced vaccine efficacy, these differences likely reflect intrinsic variations between the subclades circulating in Brazil and the vaccine strain. It is important to note that all viruses circulating in Brazil in 2021 were descendants of the 2a.2 subclade, whereas A/Hong Kong/2671/2019 belongs to the 3C.2a1b.1b subclade (1b). Nonetheless, ferret antisera raised against the vaccine strain exhibited poor recognition of all 2a.2 viruses, with titers <40 (WHO 2021b).

The discrepancy between circulating viruses and the 2021 vaccine strain prompted the update of the H3N2 component for the

subsequent season (WHO 2021b), which resulted in a reduction of mismatches (5–7) when compared to the viruses circulating in Brazil in 2022. Ferret antisera raised against the egg-propagated A/Darwin/9/2021 strain exhibited strong reactivity against 2a.2 viruses, particularly those containing the HA1 substitution H156S (WHO 2021b, 2022b). In contrast, viruses from subclades 2b and 2a.1b exhibited reduced reactivity to this antiserum (WHO 2023b). Notably, viruses from both subclades circulating in Brazil harbored a different residue at position 156: S156H. By the end of 2022, viruses belonging to the 2b subclade became predominant in Brazil. It is important to note that a lower number of mismatches does not necessarily correlate with effective vaccine coverage.

Beyond this mutation, one (G225D) is shared among the dominant subclades, and four others (E50K, F79V, I140K, and G186D) are typically encoded by viruses from the 2b subclade (WHO 2023b). Notably, the G225D, I140K, and N186D are in the sialic acid binding site, and the E50K and F79V substitutions are situated in the F′ and VE domains (Bangaru et al. 2018, Zheng et al. 2018, Wu and Wilson 2020). Mutations within the RBD can alter the interaction between the HA protein and alpha-2,3 or alpha-2,6 sialic acids across diverse host species (Zhao and Pu 2022). Despite its general conservation, the receptor-binding site is prone to selective pressures, novel mutations, and antibody evasion (Allen and Ross 2018). The G225D substitution resulted in a charge change from an uncharged polar residue, potentially allowing circulating strains to evade recognition by vaccine-induced antibodies. The F′ domain, located within the stem domain and implicated as a site of antibody binding, facilitates low pH-triggered membrane fusion activity of HA within endosomal compartments (Sriwilaijaroen and Suzuki 2012, Mair et al. 2014, Bangaru et al. 2018, Wu and Wilson 2020). The amino acid substitutions in the F′ domain, particularly those altering its electrostatic profile at lower pH levels, play a pivotal role in modifying viral infection dynamics (Bangaru et al. 2018). Regarding VE domain mutations, while its function remains unclear for influenza A and B viruses, it is known to be crucial for interaction with monoclonal antibodies, and alterations in this region could impact viral antigenicity (Crowe 2019). Indeed, the postinfection ferret antisera raised against the egg-propagated A/Darwin/9/2021-like viruses exhibited a less favorable reaction (titers = 80), especially toward viruses expressing 2a.3a.1 or 2b HA genes (WHO 2023a). Conversely, ferret antisera raised against 2a.3a.1-like viruses (e.g. egg-propagated A/Thailand/8/2022, season SH 2024) demonstrated robust recognition of most circulating viruses (WHO 2023a) and suggests a closer similarity between the circulating strains and the vaccine strain.

Considering the panorama of the influenza B virus in Brazil, the literature on the subject is limited. Therefore, our investigation represents an essential study in advancing our understanding of this virus in the country. Our analysis revealed that a single main clade of influenza B, identified as V1A.3a.2, has circulated in Brazil. This clade has been exclusively circulating not only in Brazil but also in other regions across the globe since February 2023, with a noteworthy substitution of the Yamagata lineage (da Silva et al. 2022, Kissling et al. 2023, Sominina et al. 2023, Zhou et al. 2023). Our findings are consistent with previous studies conducted in the North and Northeast regions of Brazil, highlighting the unique dominance of the V1A.3a.2 clade in these areas (Das Chagas Junior et al. 2023). The ongoing diversification of the clade has resulted in the emergence and near fixation of distinct subclades, some of which exhibit recurrent HA1:197E and HA1:183K substitutions (ECDC 2024, Huddleston et al. 2024). Recent viruses carrying the D197E mutation have been preliminarily classified as the C.5 clade, with seven descending major subclades (C.5.1–C.5.7) identified globally (ECDC 2024). In this study, four subclades (C.5.1, C.5.2, C.5.3, and C.5.4) were identified as circulating in Brazil. Most sequences were recovered in C.5, followed by subclades C.5.3 and C.5.2. According to our phylogenetic analysis, clusters VII and IX are likely to be assigned to clade C.5, while clusters V and VI may be designated as C.5.3 and C.5.2, respectively. Both haplotypes C.5.2 and C.5.3 have been predominantly recorded in South America, suggesting their probable origin in this region (available at https://nextstrain.org/flu/seasonal/vic/ha/2y, accessed 5 April 2024).

The amino acid substitutions D197E and Q200P were identified in the consensus sequences of circulating Brazilian viruses in 2022. However, only the D197E substitution was observed in 2023. The WHO has previously documented the D197E mutation, noting that its presence does not indicate a decrease in vaccine efficacy (WHO 2023a). Furthermore, none of the sites identified as being under positive selective pressure in the clade suffered modification compared to the vaccine strain. This finding is consistent with integrated hemagglutination inhibition data and molecular evolution analyses of the HA segment, which suggest effective coverage against recent V1A.3a.2 strains by the current vaccine strain B/Austria/1359417/2021 (WHO 2021b, 2022b, 2023b, Huddleston et al. 2024).

The analysis of viral alternation patterns revealed in this study provides a detailed insight into the circulation of influenza viruses in Brazil. This dynamic closely aligns with the epidemiological data provided by the Brazilian Ministry of Health for the specified period (BRASIL 2022, 2023b). A similar trend of low A/H3N2 HA sequence counts and an increase of both A/H1N1pdm09 and B/Victoria circulation were observed across different global regions throughout 2022, including Europe, North America, and Oceania (Chen et al. 2022, WHO 2022b, ECDC 2023a). Furthermore, between February and September 2023, a notable increase in the detection of the B/Victoria virus, from 5.8% to approximately one-third of globally detected viruses, was observed (WHO 2023a, 2023b).

Based on the data regarding the circulation of influenza, we determined distinct patterns of circulation across Brazilian regions. This can be explained by the continental size of the country, which encompasses different climatic zones and demographics. In fact, it has been observed that countries in tropical and subtropical zones can display influenza multiple peaks and sustained activity, often correlating with the rainy season (Moura 2010, Tamerius et al. 2013). Factors such as climate, location, and road infrastructure likely influence influenza seasonality and spread dynamics (Almeida et al. 2018). Notably, these factors contribute to the distinct circulation patterns observed in different regions, emphasizing the need for tailored surveillance and vaccination strategies to effectively manage influenza outbreaks nationwide. Additionally, the Brazilian Ministry of Health's decision to schedule vaccination campaigns based on regional climate variations (BRASIL 2024) further underscores the importance of such considerations in public health planning and response efforts.

In summary, our investigation unveiled significant findings, including the detection of clades exhibiting mutations suggestive of antigenic drift, heightened diversity within the A/H3N2 subtype, and distinct distribution patterns across different regions of Brazil. This research highlighted the critical importance of monitoring and understanding the molecular epidemiology of influenza viruses. Such insights are instrumental in guiding

public health agencies to implement effective measures for mitigating transmission risks and controlling outbreaks. Furthermore, the knowledge gained from this study can contribute to the development of vaccine strategies for a continental country.

Acknowledgements

The authors are grateful to Durval de Moraes Junior and Claudia Anania Santos da Silva for, respectively, logistic and information technology support.

Supplementary data

Supplementary data is available at *VEVOLU Journal* online.

Conflict of interest: None declared.

Funding

This project was supported by the São Paulo Research Foundation—FAPESP (grant no. 2021/11944-6, and fellowships for I.C.B. 2023/07688-0 and V.C.d.S. 2023/11920-5) and Fundação Butantan. S.K., E.A., E.G.K., S.N.S., S.C.S., and M.C.E. are fellows from the National Council for Scientific and Technological Development (CNPq: 306199/2018-1).

Data availability

All of the sequences generated in this study were deposited in GISAID and are publically available. Their accession numbers are presented in Supplementary Table S1.

References

Aksamentov I, Roemer C, Hodcroft E *et al*. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *J Open Source Softw* 2021;**6**:3773.

Allen JD, Ross TM. H3N2 influenza viruses in humans: viral mechanisms, evolution, and evaluation. *Hum Vaccines Immunother* 2018;**14**:1840–47.

Almeida A, Codeço C, Luz P. Seasonal dynamics of influenza in Brazil: the latitude effect. *BMC Infect Dis* 2018;**18**:1–9.

Attwood SW, Hill SC, Aanensen DM *et al*. Phylogenetic and phylodynamic approaches to understanding and combating the early SARS-CoV-2 pandemic. *Nat Rev Genet* 2022;**23**:547–62.

Bangaru S, Zhang H, Gilchuk IM *et al*. A multifunctional human monoclonal neutralizing antibody that targets a unique conserved epitope on influenza HA. *Nat Commun* 2018;**9**:2669.

Bankevich A, Nurk S, Antipov D *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;**19**:455–77.

Bedford T, Suchard MA, Lemey P *et al*. Integrating influenza antigenic dynamics with molecular evolution. *eLife* 2014;**3**:1–26.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20.

Borges V, Pinheiro M, Pechirra P *et al*. INSaFLU: an automated open web-based bioinformatics suite "from-reads" for influenza whole-genome-sequencing-based surveillance. *Genome Med* 2018;**10**:46.

BRASIL. *Informe Técnico 23ª Campanha Nacional de Vacinação Contra a Influenza Brasília, 2021, Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde (SVS) Departamento de Imunização e Vigilância de Doenças Transmissíveis (DEIDT) Coordenação Geral do Programa Nacional de Imunizações (CGPNI)*. Brasil, Ministério da Saúde, Secretaria de Vigilância em Saúde (SVS), Departamento de Imunização e Vigilância de Doenças Transmissíveis (DEIDT), Coordenação Geral do Programa Nacional de Imunizações (CGPNI). 2021a. https://www.gov.br/saude/pt-br/media/pdf/2021/marco/16/informe-tecnico-influenza-2021.pdf (20 October 2024, date last accessed).

BRASIL. *Vigilância Genômica do vírus SARS-CoV-2 no âmbito da SVS/MS*. Brasília, DF. 2021b. http://bvsms.saude.gov.br/bvs/publicacoes/vigilancia_genomica_SARS-CoV-2_ambito_SVS.pdf (10 May 2024, date last accessed).

BRASIL. *Informe Técnico: 24ª Campanha Nacional de Vacinação contra a Influenza (Versão Atualizada)*. Brasília, DF. 2022. https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/c/calendario-nacional-de-vacinacao/arquivos/informe-da-24a-campanha-nacional-de-vacinacao-contra-a-influenza.pdf (20 October 2024, date last accessed).

BRASIL. *INFORME TÉCNICO OPERACIONAL Vacinação contra a Influenza*. Brasília, DF. 2023a. https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/c/calendario-nacional-de-vacinacao/arquivos/informe-tecnico-operacional-de-vacinacao-contra-a-influenza-2023/view (05 August 2024, date last accessed).

BRASIL. *Informe Vigilância das Síndromes Gripais: Influenza, Covid-19 e outros vírus respiratórios de importância em Saúde Pública - Semana epidemiológica 31*. Brasília, DF. 2023b. https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/g/gripe-influenza/situacao-epidemiologica/informe-epidemiologico-de-influenza-covid-19-e-outros-virus-respiratorios-de-importancia-em-saude-publica-se-31 (22 April 2024, date last accessed).

BRASIL. *Estratégia de Vacinação contra a Influenza Regiões Nordeste, Centro-Oeste, Sul, e Sudeste*. 2024. https://www.gov.br/saude/pt-br/vacinacao/informes-tecnicos/estrategia-de-vacinacao-influenza-2024/view (5 August 2024, date last accessed).

Bush RM, Bender CA, Subbarao K *et al*. Predicting the evolution of human influenza A. *Science* 1999;**286**:1921–25.

Cantarino L, Merchan-Hamann E. Influenza in Brazil: surveillance pathways. *J Infect Dev Ctries* 2016;**10**:13–23.

Chan J, Holmes A, Rabadan R. Network analysis of global influenza spread. *PLoS Comput Biol* 2010;**6**:e1001005.

Chen R, Holmes EC. The evolutionary dynamics of human influenza B virus. *J Mol Evol* 2008;**66**:655–63.

Chen Z *et al*. Antigenic drift and epidemiological severity of seasonal influenza in Canada. *Sci Rep* 2022;**12**:15625.

Costa JCD, Siqueira MM, Brown D *et al*. Vaccine mismatches, viral circulation, and clinical severity patterns of influenza B Victoria and Yamagata infections in Brazil over the decade 2010–2020: a statistical and phylogeny–trait analyses. *Viruses* 2022;**14**:1477.

Crowe JE. Antibody determinants of influenza immunity. *J Infect Dis* 2019;**219**:S21–9.

Das Chagas Junior WD, Silva AM, Barbagelata LS *et al*. Predominância da linhagem Victoria do vírus influenza B durante a temporada de influenza 2023 nas regiões norte e nordeste do Brasil. *Braz J Infect Dis* 2023;**27**:103480.

da Silva DBB, de Oliveira Santos KC, Benega MA *et al*. Differentiation of influenza B lineages circulating in different regions of Brazil, 2014–2016, using molecular assay. *Vaccine: X* 2022;**12**:100220.

Dhanasekaran V, Sullivan S, Edwards KM *et al*. Human seasonal influenza under COVID-19 and the potential consequences of influenza lineage elimination. *Nat Commun* 2022;**13**:1721.

Dolinsky TJ, Nielsen JE, McCammon JA *et al*. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* 2004;**32**:W665–W667.

ECDC. Influenza virus characterization: summary report, Europe, August 2023. Copenhagen, 2023a.

ECDC. Influenza virus characterization: summary report, Europe, February 2023. Copenhagen and Stockholm, 2023b.

ECDC. *Influenza Virus Characterization: Summary Report, Europe, December 2023*. Copenhagen. 2024. https://www.who.int/europe/publications/i/item/WHO-EURO-2023-6189-45954-72777 (5 April 2024, date last accessed).

Ferguson NM, Galvani AP, Bush RM. Ecological and immunological determinants of influenza evolution. *Nature* 2003;**422**:428–33.

Fitch WM, Bush RM, Bender CA *et al*. Long term trends in the evolution of H(3) HA1 human influenza type A. *Proc Natl Acad Sci* 1997;**94**:7712–18.

Forleo-Neto E, Halker E, Santos VJ *et al*. Influenza. *Rev Soc Bras Med Trop* 2003;**36**:267–74.

Glezen PW, Schmier JK, Kuehn CM *et al*. The burden of influenza B: a structured literature review. *Am J Public Health* 2013;**103**:e43–e51.

Grubaugh ND, Gangavarapu K, Quick J *et al*. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol* 2019;**20**:8.

Hadfield J, Megill C, Bell SM *et al*. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 2018;**34**:4121–23.

Huddleston J, Bedford T, Chang J *et al*. 2024. *Seasonal influenza circulation patterns and projections for February 2024 to February 2025*. https://bedford.io/papers/huddleston-feb-2024-flu-report/ (15 May 2023, date last accessed).

Irving SA, Patel DC, Kieke BA *et al*. Comparison of clinical features and outcomes of medically attended influenza A and influenza B in a defined population over four seasons: 2004–2005 through 2007–2008. *Influenza Other Respir Viruses* 2012;**6**:37–43.

Jurrus E, Engel D, Star K *et al*. Improvements to the APBS biomolecular solvation software suite. *Protein Sci* 2018;**27**:112–28.

Kalyaanamoorthy S, Minh BQ, Wong TKF *et al*. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;**14**:587–89.

Keech M, Beardsworth P. The impact of influenza on working days lost. *PharmacoEconomics* 2008;**26**:911–24.

Kissling E, Pozo F, Martínez-Baz I *et al*. Influenza vaccine effectiveness against influenza A subtypes in Europe: results from the 2021–2022 I-MOVE primary care multicentre study. *Influenza Other Respir Viruses* 2023;**17**:e13069.

Kosakovsky Pond SL, Frost SDW, Muse SV *et al*. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 2005;**21**:676–79.

Koutsakos M, Nguyen TH, Barclay WS *et al*. Knowns and unknowns of influenza B viruses. *Future Microbiol* 2016;**11**:119–35.

Ladner JT, Sahl JW. Towards a post-pandemic future for global pathogen genome sequencing. *PLoS Biol* 2023;**21**:e3002225.

Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–59.

Li H, Birol I. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 2018;**34**:3094–100.

Mair CM, Ludwig K, Herrmann A *et al*. Receptor binding and pH stability—how influenza A virus hemagglutinin affects host-specific virus infection. *Biochim Biophys Acta Biomembr* 2014;**1838**:1153–68.

McCullers JA, Hayden FG. Fatal influenza B infections: time to reexamine influenza research priorities. *J Infect Dis* 2012;**205**:870–72.

Mena I, Nelson MI, Quezada-Monroy F *et al*. Origins of the 2009 H1N1 influenza pandemic in swine in Mexico. *eLife* 2016;**5**:1–21.

Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* 2013;**30**:1188–95.

Moura FE. Influenza in the tropics. *Curr Opin Infect Dis* 2010;**23**:415–20.

Murrell B, Moola S, Mabona A *et al*. FUBAR: A fast, unconstrained Bayesian AppRoximation for inferring selection. *Mol Biol Evol* 2013;**30**:1196–205.

Nair H, Brooks WA, Katz M *et al*. Global burden of respiratory infections due to seasonal influenza in young children: a systematic review and meta-analysis. *Lancet* 2011;**378**:1917–30.

Nguyen L-T, Schmidt HA, von Haeseler A *et al*. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268–74.

Olson DR, Heffernan RT, Paladini M *et al*. Monitoring the impact of influenza by age: emergency department fever and respiratory complaint surveillance in New York City. *PLoS Med* 2007;**4**:1349–61.

Olsson MHM, Søndergaard CR, Rostkowski M *et al*. PROPKA3: consistent treatment of internal and surface residues in empirical pK_a predictions. *J Chem Theory Comput* 2011;**7**:525–37.

Paget J, Spreeuwenberg P, Charu V *et al*. Global mortality associated with seasonal influenza epidemics: new burden estimates and predictors from the GLaMOR Project. *J Glob Health* 2019;**9**:1–12.

Petrova VN, Russell CA. The evolution of seasonal influenza viruses. *Nat Rev Microbiol* 2018;**16**:47–60.

Pillai TK, Johnson KE, Song T *et al*. Tracking the emergence of antigenic variants in influenza A virus epidemics in Brazil. *Virus Evol* 2023;**9**:1–7.

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;**26**:841–42.

Rana R, Singhal R. Chi-square test and its application in hypothesis testing. *J Pract Cardiovasc Sci* 2015;**1**:69.

Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 1993;**234**:779–815.

Sharma Y, Horwood C, Hakendorf P *et al*. Clinical characteristics and outcomes of influenza A and B virus infection in adult Australian hospitalised patients. *BMC Infect Dis* 2020;**20**:1–9.

Shen W, Le S, Li Y *et al*. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 2016;**11**:e0163962.

Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data—from vision to reality. *Euro Surveill* 2017;**22**:30494.

Smith DJ, Lapedes AS, de Jong JC *et al*. Mapping the antigenic and genetic evolution of influenza virus. *Science* 2004;**305**:371–76.

Sominina A, Danilenko D, Komissarov AB *et al*. Assessing the intense influenza A(H1N1)pdm09 epidemic and vaccine effectiveness in the post-COVID season in the Russian Federation. *Viruses* 2023;**15**:1780.

Søndergaard CR, Olsson MHM, Rostkowski M *et al*. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pK_a values. *J Chem Theory Comput* 2011;**7**:2284–95.

Southgate JA, Bull MJ, Brown CM *et al*. Influenza classification from short reads with VAPOR facilitates robust mapping pipelines and zoonotic strain detection for routine surveillance applications. *Bioinformatics* 2020;**36**:1681–88.

Sriwilaijaroen N, Suzuki Y. Molecular basis of the structure and function of H1 hemagglutinin of influenza virus. *Proc Jpn Acad Ser B* 2012;**88**:226–49.

Tamerius JD, Shaman J, Alonso WJ *et al*. Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLoS Pathog* 2013;**9**:e1003194.

Tapia R, Torremorell M, Culhane M *et al*. Antigenic characterization of novel H1 influenza A viruses in swine. *Sci Rep* 2020;**10**:4510.

Thompson WW, Weintraub E, Dhankhar P *et al*. Estimates of US influenza-associated deaths made using four different methods. *Influenza Other Respir Viruses* 2009;**3**:37–49.

Troeger CE, Blacker BF, Khalil IA *et al*. Mortality, morbidity, and hospitalisations due to influenza lower respiratory tract infections, 2017: an analysis for the Global Burden of Disease Study 2017. *Lancet Respir Med* 2019;**7**:69–89.

Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010;**31**:455–61.

WHO. *Recommended Composition of Influenza Virus Vaccines for Use in the 2020–2021 Northern Hemisphere Influenza Season*. 2020a. http://www.who.int/influenza/vaccines/virus/en/ (7 November 2023, date last accessed).

WHO. *Recommended Composition of Influenza Virus Vaccines for Use in the 2021 Southern Hemisphere Influenza Season*. 2020b. http://www.who.int/influenza/vaccines/virus/en (20 December 2023, date last accessed).

WHO. *Recommended Composition of Influenza Virus Vaccines for Use in the 2021–2022 Northern Hemisphere Influenza Season*. 2021a. http://www.who.int/influenza/vaccines/virus/en/ (20 December 2023, date last accessed).

WHO. *Recommended Composition of Influenza Virus Vaccines for Use in the 2022 Southern Hemisphere Influenza Season*. 2021b. https://www.who.int/teams/global-influenza-programme/vaccines/vaccine-in-tropics-and-subtropics (6 April 2023, date last accessed).

WHO. *Recommended Composition of Influenza Virus Vaccines for Use in the 2022–2023 Northern Hemisphere Influenza Season*. 2022a. https://www.who.int/teams/global-influenza-programme/vaccines/who-recommendations (20 December 2023, date last accessed).

WHO. *Recommended Composition of Influenza Virus Vaccines for Use in the 2023 Southern Hemisphere Influenza Season*. 2022b. https://cdn.who.int/media/docs/default-source/influenza/who-influenza-recommendations/vcm-southern-hemisphere-recommendation-2023/202209_recommendation.pdf?sfvrsn=83a26d50_3&download=true (7 November 2023, date last accessed).

WHO. *Recommended Composition of Influenza Virus Vaccines for Use in the 2023–2024 Northern Hemisphere Influenza Season*. 2023a. https://www.who.int/publications/m/item/recommended-composition-of-influenza-virus-vaccines-for-use-in-the-2024-southern-hemisphere-influenza-season (20 December 2023, date last accessed).

WHO. *Recommended Composition of Influenza Virus Vaccines for Use in the 2024 Southern Hemisphere Influenza Season*. 2023b. https://www.who.int/publications/m/item/recommended-composition-of-influenza-virus-vaccines-for-use-in-the-2024-southern-hemisphere-influenza-season (20 December 2023, date last accessed).

Wu NC, Wilson IA. Influenza hemagglutinin structures and antibody recognition. *Cold Spring Harb Perspect Med* 2020;**10**:a038778.

Yamayoshi S, Kawaoka Y. Current and future influenza vaccines. *Nature Med* 2019;**25**:212–20.

Yu G, Smith DK, Zhu H *et al*. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Meth Ecol Evolut* 2017;**8**:28–36.

Zhao C, Pu J. Influence of host sialic acid receptors structure on the host specificity of influenza viruses. *Viruses* 2022;**14**:2141.

Zheng Z, Paul SS, Mo X *et al*. The vestigial esterase domain of haemagglutinin of H5N1 avian influenza A virus: antigenicity and contribution to viral pathogenesis. *Vaccines* 2018;**6**:53.

Zhou B, Lin X, Wang W *et al*. Universal influenza B virus genomic amplification facilitates sequencing, diagnostics, and reverse genetics. *J Clin Microbiol* 2014;**52**:1330–37.

Zhou X, Lin Z, Tu J *et al*. Persistent predominance of the Victoria lineage of influenza B virus during COVID-19 epidemic in Nanchang, China. *Influenza Other Respir Viruses* 2023;**17**:e13226.

Zolotarova O, Fesenko A, Holubka O *et al*. Genotypic variants of pandemic H1N1 influenza A viruses isolated from severe acute respiratory infections in Ukraine during the 2015/16 influenza season. *Viruses* 2021;**13**:2125.

Virus Evolution, 2025, 11(1), veae102 , DOI: https://doi.org/10.1093/ve/veae102, Advance Access Publication Date: 23 November 2024, Research Article