

Emergence and predominance of novel influenza A (H3N2) subclade (3C.2a1b.2a.2a.3a.1) in Kenya and globally in 2023

John Mwita Morobe, Edidah Moraa, Arnold W. Lambisia, Martin Mutunga, Brenda Kiage, Esther Nyadzua Katama, Timothy Makori, Joyce Nyiro, Leonard Ndwiga, James Nyagwange, Charles Sande, Lynette Isabella Ochola-Oyier, Edward C. Holmes, D. James Nokes, George Githinji, Charles N. Agoti

Author affiliations: Kenya Medical Research Institute Wellcome Trust Research Programme, Kilifi, Kenya (John Mwita Morobe, Edidah Moraa, Arnold W. Lambisia, Timothy Makori, Joyce Nyiro, Leonard Ndwiga, James Nyagwange, Isabella Oyier, James Nokes, George Githinji, Charles N. Agoti); The University of Sydney (Edward Holmes); University of Warwick, Coventry, UK (James Nokes); Pwani University, Kilifi (Charles N. Agoti, G. Githinji)

Abstract

We report emergence and predominance of the influenza A (H3N2) subclade 3C.2a1b.2a.2a.3a.1 in Kenya similar to the global clade in 2023. The Kenyan 3C.2a1b.2a.2a.3a.1 viruses have >15 amino acid differences in the HA and NA proteins relative to 2023/24 WHO recommended Northern/Southern Hemisphere influenza vaccine strains.

Main

Antigenic drift in the hemagglutinin (HA) and neuraminidase (NA) proteins of influenza A/B viruses is a key contributor to their recurrent global epidemics¹. Influenza genomic surveillance helps identify virus strains likely to predominate in future epidemics, enabling appropriate vaccine strain recommendations². Critically, however, the limited influenza genomic surveillance in sub-Saharan Africa obscures global influenza epidemiology.

The KEMRI-Wellcome Trust Research Programme, coastal Kenya, runs two respiratory pathogen surveillance platforms: (a) paediatric inpatient surveillance of severe and very severe pneumonia in <5-year-olds at Kilifi County Hospital³ and (b) all age groups outpatient acute respiratory infection (ARI) surveillance at five Kilifi

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

County health facilities (HF)⁴. The former surveillance collects respiratory samples from all eligible consented patients, while the latter collects ~15 samples per HF per week. Influenza infection is diagnosed using RT-PCR⁵.

Between 1st January and 29th September 2023, 3,262 respiratory samples were collected from the two platforms⁶. Only 1,192 (37%) of collected samples were screened due to reagent limitations. Of these, 6% (68/1192) were positive for influenza A (n=58, 85%), B (n=6, 9%) or C (n=4, 6%) viruses; The weekly influenza A/B/C detections are shown in **Figure 1A**.

Influenza A positives were targeted for a multi-segment PCR⁷ and sequencing on Illumina MiSeq or ONT GridION. Raw reads were assembled using the reference-based FLU or FLU-MinION module of the Iterative Refinement Meta-Assembler (IRMA)⁸. Recovered sequences were assigned phylogenetic clades /subclades based on HA gene using NextClade⁹. Maximum likelihood (ML) phylogenetic trees were estimated using IQ-TREE v2.2.2.6¹⁰, and time-scaled phylogenies were estimated using TimeTree¹¹.

Full-length viral genomes (>90% coverage) were recovered from 52 of 58 samples and deposited in GISAID EpiFlu™ (accession nos EPI_ISL_18560856 - EPI_ISL_18560859 and EPI_ISL_18560907 - EPI_ISL_18560947). These were classified as A(H1N1) pdm09 (n=6, 12%) and A(H3N2) (n=46, 88%). Three clades/subclades were identified: one within A(H1N1) pdm09 and two within A(H3N2). The (H3N2) 3C.2a1b.2a.2a.3a.1 viruses, first seen in the Kenya on 19th May 2023, displaced previously circulating clades/subclades within the A(H3N2) subtype (**Figure 1B**). Based on GISAID data, the 3C.2a1b.2a.2a.3a.1 viruses were first detected in United Arab Emirates in January 2022, and have displaced other A(H3N2) clades/sub-clades globally (**Figure 1C**). The Kenya patients infected with 3C.2a1b.2a.2a.3a.1 viruses presented with cough (100%), fever (82%), nasal discharge (80%) and difficulty breathing (13%) (**Table 1**).

The global phylogeny of the 3C.2a1b.2a.2a.3a.1 subclade (n=3,462) revealed that the Kenyan sequences fall into 4 groups (two clusters, named C1 and C2, and 2 singletons, named S1 and S2) indicating at least 4 separate introductions into the

country (**Appendix 1 Figure 1**). The Kenya HA and NA sequences differed from the global sequences at multiple nucleotide positions: NA (G435A, A687G, A1230G and C1275T) and HA (deletion at 131, G991A, A1035G and G1124A). Based on the HA gene segment, the mean time to the most recent common ancestor for the 3C.2a1b.2a.2a.3a.1 subclade viruses was estimated to be April 2019 [95% Highest Posterior Density (HPD): August 2018 – January 2020], with a nucleotide substitution rate of 1.19×10^{-3} [95% HPD: $9.53 \times 10^{-4} - 1.40 \times 10^{-3}$] substitutions/site/year.

There were amino acid changes between the HA of the Kenyan 3C.2a1b.2a.2a.3a.1 viruses and the 2023/24 season WHO recommended vaccine strains: 26 and 27 amino acid changes relative to the HA protein of the Northern Hemisphere (NH) vaccine strains – A/Darwin/6/2021 (H3N2)-like virus and A/Darwin/9/2021(H3N2)-like virus, respectively (**Appendix 1 Figure 2A and B**). The substitutions included T3A (signal peptide), E50K, D53N, G53N, N96S, I140K, N186D, I192F, I223V, G225D in HA1, and N49S (HA2). These mutations were seen in 100% of the Kenya viruses. There were 17 amino acid changes relative to the HA of the Southern Hemisphere (SH) vaccine strain – A/Massachusetts/18/2022 (H3N2)-like virus, (**Appendix 1 Figure 2C**). Substitutions included I58V(HA1) observed in 11% of the viruses, V77I (HA2) observed in 8% of the viruses. Comparison of the Kenya viruses to the SH vaccine strain A/Thailand/18/2022 (H3N2)-like virus also revealed 17 amino acid changes (**Appendix 1 Figure 2D**) with the most frequent HA substitution being I58V(HA1) observed in 11% of the viruses, V77I (HA2) observed in 8% of the viruses.

Relative to the NA protein of Northern Hemisphere (NH) vaccine strains A/Darwin/6/2021 (H3N2)-like virus and A/Darwin/9/2021(H3N2)-like virus, showed 22 and 21 amino acid changes, respectively (**Appendix 1 Figure 3A and B**). A comparison with SH vaccine strain A/Thailand/18/2022 (H3N2)-like virus and A/Massachusetts/18/2022 (H3N2)-like virus revealed 21 and 20 amino acid changes respectively in NA protein (**Appendix 1 Figure 3C and D**).

The factors contributing to the global emergence and predominance of 3C.2a1b.2a.2a.3a.1 viruses over other A(H3N2) strains are currently unclear. We

speculate that the H1 mutations observed in this subclade likely enhance fitness over the other subclades^{12,13}, or make it less immunologically recognizable to current immunity established by vaccination or natural infection¹⁴. Our findings emphasize the need for near-real time genomic surveillance globally, including LMICs, to monitor the evolution of circulating IAV for optimal vaccine strain selection.

Acknowledgments

We thank the members of the Viral Epidemiology and Control Research Group at KEMRI Wellcome Trust Research Programme, field, clinical and laboratory staff, for their valuable contributions to this work. We also thank laboratories that generated and shared genetic sequence data via the GISAID Initiative (<https://www.gisaid.org>), on which this part of this research is based. Genome sequences generated in this study are available in the GISAID's EpiFlu database (accession nos EPI_ISL_18560856 - EPI_ISL_18560859 and EPI_ISL_18560907 - EPI_ISL_18560947). The dataset and analysis scripts used are available in the Harvard Dataverse repository (<https://doi.org/10.7910/DVN/5US8MM>).

Funding

This work was supported by The Wellcome, UK (grant nos. 226002/A/22/Z and 220985/Z/20/Z), New Variant Assessment Programme, a UK Health Security Agency program funded by the UK Department of Health and Social Care as a global initiative to strengthen genomic surveillance for pandemic preparedness and response to emerging and priority infectious diseases; the UK National Institute for Health and Care Research (project references 17/63/82 and 16/136/33) that uses aid from the UK government to support global health research; UK Foreign, Commonwealth and Development Office. The views expressed in this publication are those of the author(s) and not necessarily those of Wellcome or National Institute for Health and Care Research, Department of Health and Social Care, or the Foreign Commonwealth and Development Office.

Reference

1. Petrova VN, Russell CA. The evolution of seasonal influenza viruses. *Nat Rev Microbiol.* 2018;16(1):47-60. doi:10.1038/nrmicro.2017.118

2. McHardy AC, Adams B. The Role of Genomics in Tracking the Evolution of Influenza A Virus. *PLoS Pathog.* 2009;5(10):e1000566-.
<https://doi.org/10.1371/journal.ppat.1000566>
3. Nokes DJ, Ngama M, Bett A, et al. Incidence and severity of respiratory syncytial virus pneumonia in rural Kenyan children identified through hospital surveillance. *Clin Infect Dis.* 2009;49(9):1341-1349. doi:10.1086/606055
4. Scott JAG, Bauni E, Moisi JC, et al. Profile: The Kilifi Health and Demographic Surveillance System (KHDSS). *Int J Epidemiol.* 2012;41(3):650-657.
doi:10.1093/ije/dys062
5. Gunson RN, Collins TC, Carman WF. Real-time RT-PCR detection of 12 respiratory viral infections in four triplex reactions. *J Clin Virol.* 2005;33(4):341-344. doi:10.1016/j.jcv.2004.11.025
6. Hammit LL, Kazungu S, Morpeth SC, et al. A preliminary study of pneumonia etiology among hospitalized children in Kenya. *Clin Infect Dis.* 2012;54 Suppl 2(Suppl 2):S190-9. doi:10.1093/cid/cir1071
7. Bin Z, Yi-Mo D, R B J, et al. Multiplex Reverse Transcription-PCR for Simultaneous Surveillance of Influenza A and B Viruses. *J Clin Microbiol.* 2017;55(12):3492-3501. doi:10.1128/jcm.00957-17
8. Shepard SS, Meno S, Bahl J, Wilson MM, Barnes J, Neuhaus E. Viral deep sequencing needs an adaptive approach: IRMA, the iterative refinement meta-assembler. *BMC Genomics.* 2016;17(1):708. doi:10.1186/s12864-016-3030-6
9. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics.* 2018;34(23):4121-4123.
doi:10.1093/bioinformatics/bty407
10. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol.* 2015;32(1):268-274. doi:10.1093/molbev/msu300
11. Sagulenko P, Puller V, Neher RA. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol.* 2018;4(1):vex042-vex042.
doi:10.1093/ve/vex042
12. 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(1):1721. doi:10.1038/s41467-022-29402-5
13. Wang Y, Tang CY, Wan XF. Antigenic characterization of influenza and SARS-CoV-2 viruses. *Anal Bioanal Chem.* 2022;414(9):2841-2881.
doi:10.1007/s00216-021-03806-6
14. Kim K, Gouma S, Vieira MC, Weirick ME, Hensley SE, Cobey S. Measures of population immunity can predict the dominant clade of influenza A (H3N2) and reveal age-associated differences in susceptibility and specificity. *medRxiv.* Published online January 1, 2023:2023.10.26.23297569.
doi:10.1101/2023.10.26.23297569

Figure and Table

Table 1. Signs and symptoms present among patients infected with influenza A (H3N2) 3C.2a1b.2a.2a.3a.1 viruses, Kilifi, Kenya, January–September 2023.

Signs/symptoms	No. (%) patients (n=45)
Fever	
Yes	37 (82%)
Cough	
Yes	45 (100%)
Nasal discharge	
Yes	36 (80%)
Difficulty breathing	
Yes	6 (13%)
Sore throat	
Yes	3 (7%)
Body malaise	
Yes	10 (22%)
Chest indrawing	
Yes	0 (0%)
Wheezing	
Yes	1 (2%)

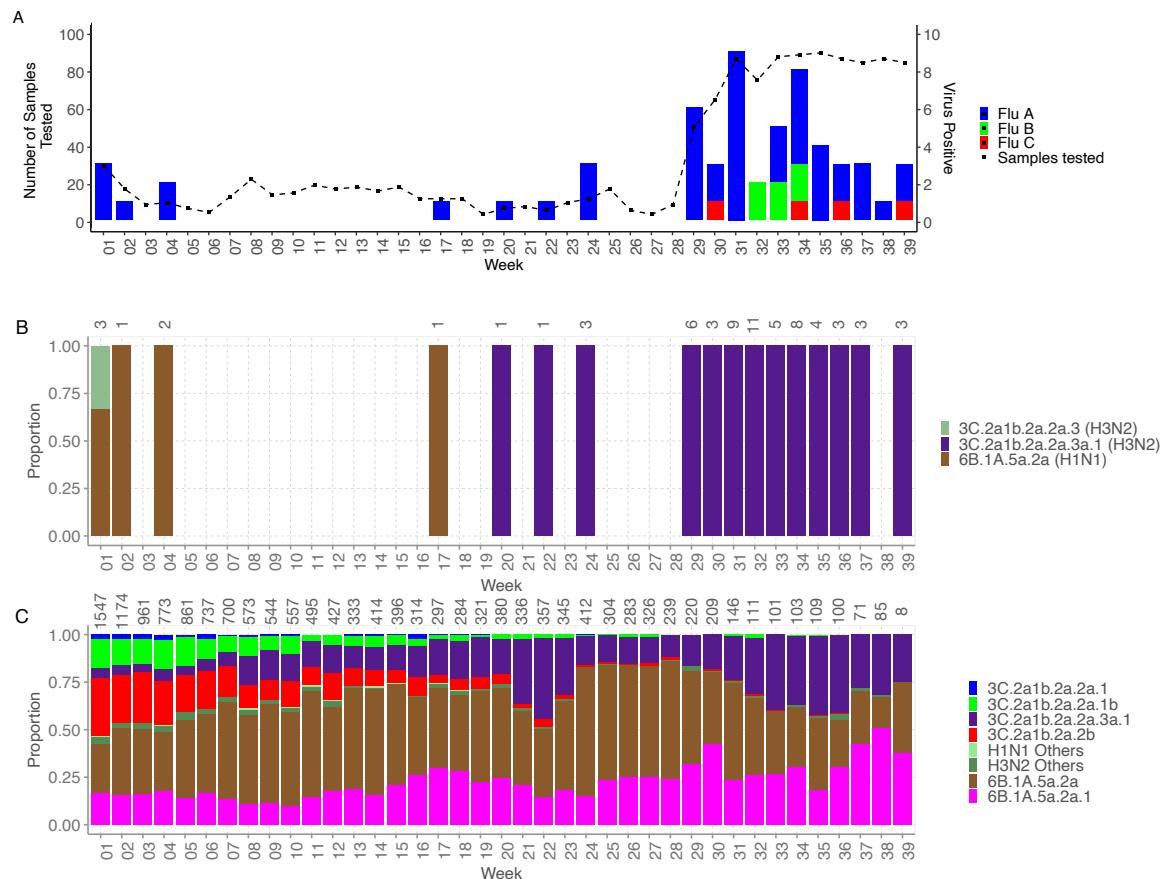
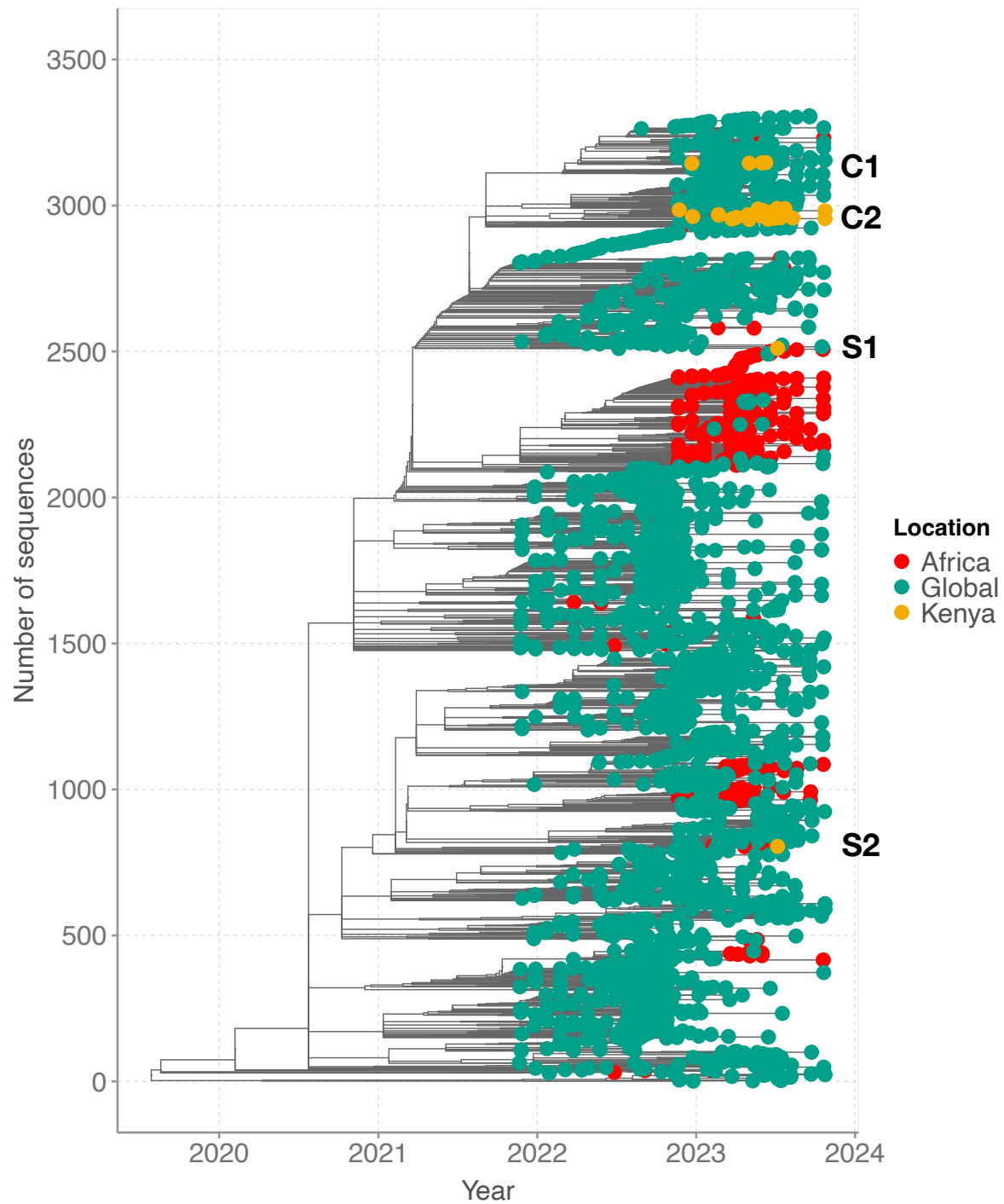
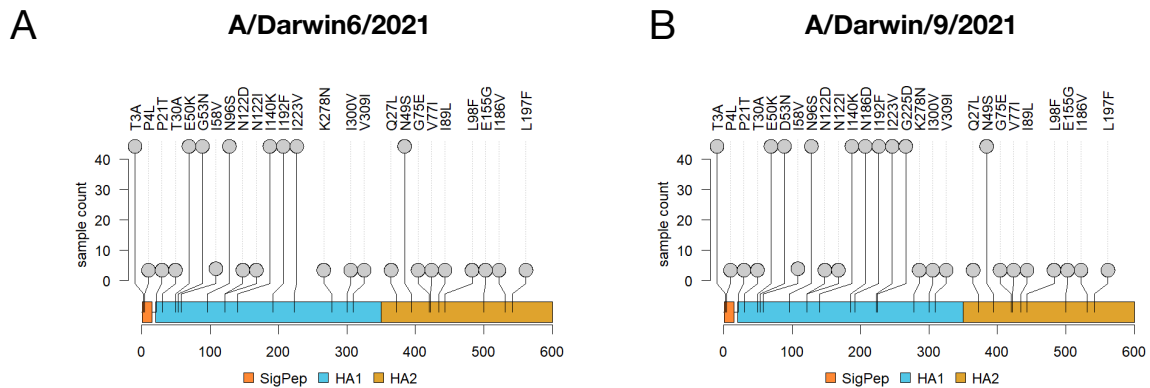


Figure 1. Temporal distribution of influenza in Kenya and globally. A) Weekly detections of influenza virus type A, B and C in samples collected in Kilifi between 1st January and 29th September 2023. The observed low testing numbers (10-20 samples per week) up to week 28, followed by a rapid increase to 80 per week is attributed to a batch of samples that were not tested during weeks 0 to 28 due to reagents constraints. B) Distribution of clades/subclades within the influenza subtype A(H3N2) pdm09 and A(H1N1) subtype in Kilifi. The numbers on top of the bars indicate the number of sequences available for each analyzed week. C) Distribution of clades/subclades within the influenza subtype A(H3N2) pdm09 and A(H1N1) subtype worldwide in 2023 (Jan-Sep) based on data deposited on GISAID (accessed on 7th November 2023).

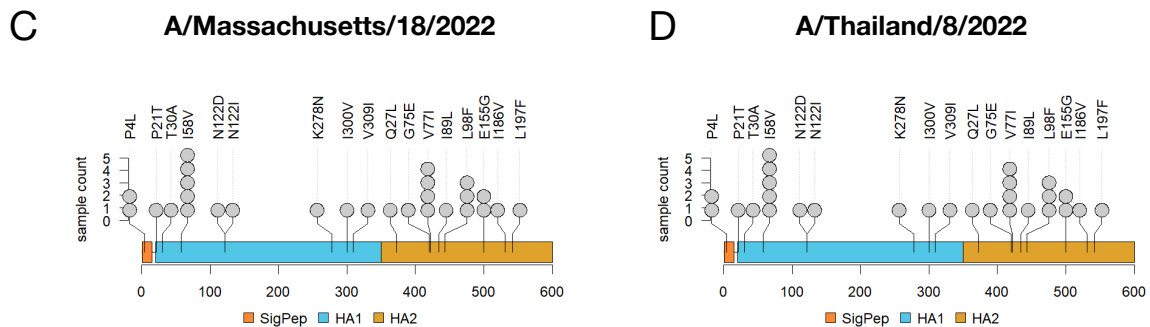


Appendix 1 Figure 1. Time-resolved phylogeny of global genomes ($n=3,462$) combined with Kilifi IAV genomes ($n=45$). Global genomes are depicted in green, African genomes in red, and Kenyan genomes in orange.

Northern Hemisphere vaccine strains

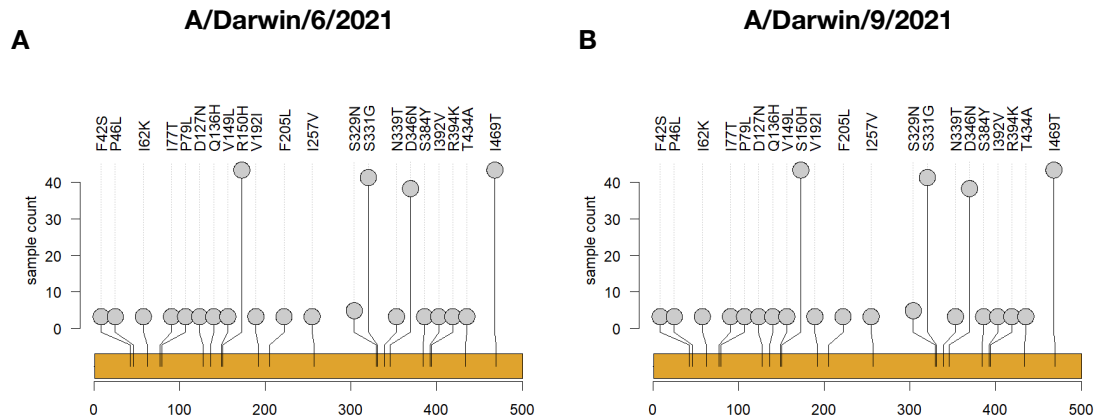


Southern Hemisphere vaccine strains

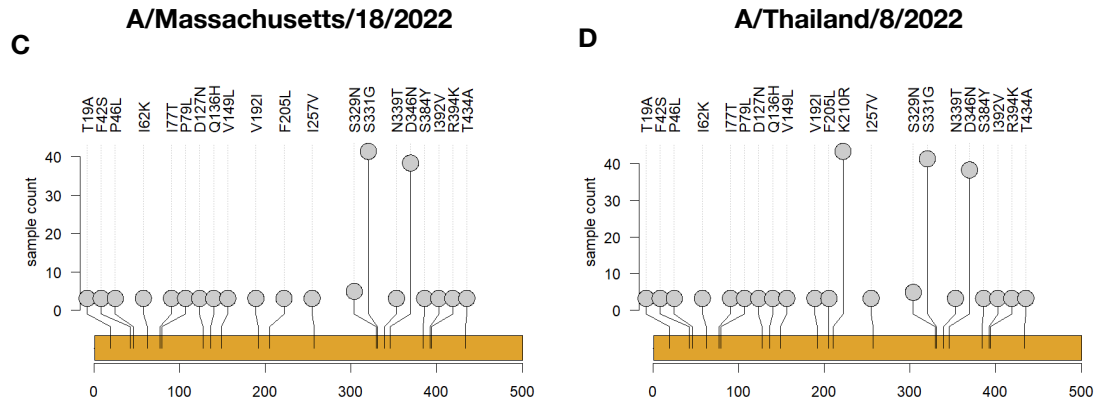


Appendix 1 Figure 2. Amino acid sequence differences in the HA protein. The lollipop plot represents the positions and amino acid difference in the HA protein in comparison to the respective Northern Hemisphere (NH) and Southern Hemisphere (SH) WHO recommended vaccine strains.

Northern Hemisphere vaccine strains



Southern Hemisphere vaccine strains



Appendix 1 Figure 3. Amino acid sequence differences in the NA protein. The lollipop plot represents the positions and amino acid difference in the NA protein in comparison to the respective Northern Hemisphere (NH) and Southern Hemisphere (SH) WHO recommended vaccine strains.