medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

PHYLOGEOGRPHY REVEALS EXPANSION OF YELLOW FEVER VIRUS GENOTYPES IN WEST AND CENTRAL AFRICA

Babatunde Olanrewaju Motayo[1*], Adewale Opayele[2], Paul Akiniyi Akinduti[3], Adedayo Omotayo Faneye[2]

1. Department of Medical Microbiology, Federal Medical Centre, Abeokuta, Nigeria.
2. Department of Virology, College of Medicine, University of Ibadan, Nigeria.
3. Department of Biological Sciences, Covenant University, Ota, Nigeria

Correspondence*: Babatunde O Motayo e-mail babatundemotayo@yahoo.com

ABSTRACT

Yellow Fever virus remains the most medically significant Arbovirus in Africa, with the occurrence of large outbreaks with human fatalities in recent years in Africa. Molecular epidemiology has shown the presence of 4 genotypes circulating across Africa, however still paucity of data still exists, regarding directional spread and phylogeography of the African Yellow fever genotypes. The need to fill in this gap with information from recent spatiotemporal events from continuous occurrence of YF outbreaks in Africa conceptualized this study; which aims to investigate the most recent transmission events and directional spread of YF virus using updated genomic sequence data. Archived Yellow Fever sequence data was utilized along with epidemiologic data from outbreaks occurring between the years to analyze the case/fatality distribution over the study period. Phylogeograhic analysis was also utilized to demonstrate the ancestral introduction and geographic clustering of YF genotypes in Africa. Directional spread and geographic transmission of YF was also investigated. African YF genotypes were found to be geographical distinct, circulating within these distinct geographical boundaries. Spatiotemporal spread however revealed a westward spread of the West African genotypes over time, and recent northward movement of the East African genotype. We conclude by recommending expanded human/ vectoral surveillance of YF and other Arboviruses of public health importance, and upscaling sequencing capabilities of new and existing public health labs in Africa to help in the defense against public health threats.

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

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

INTRODUCTION

Yellow fever virus is the causative agent for yellow fever, a viral illness that affects mainly the liver, and may cause serious complications including hemorrhage and death (Monath, 2015). It is a member of the family Flaviviridea, genus flavavirus, along with Dengue virus, West Nile virus and Japanese encephalitis virus. It is an enveloped virus with a single stranded positive-sense RNA genome of 11kb size, that encodes a single polyprotein that is cleaved into 3 structural proteins, envelope (E), matrix (M), and capsid proteins (C) and 7 non-structural proteins, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 (Beasley et al, 2015). Yellow fever is an arthropod borne infection, maintained by the mosquito vector *Aedes egypti*, and *Aedes albopticus*, although several vectorial studies have identified YF virus in other mosquitoes' species, such as Aedes *taylori* (Goud et al, 2003). The main vector of the YF virus is also capable of transmitting other viral infections such as Chikungunya (CHIK), Dengue (DENV) and Zika (ZIKV) in tropical and sub-tropical areas (Goud et al, 2003). There are two major vectorial cycles of YF, the sylvatic also known as Jungle YF maintained in non-human primates (NHP) by the mosquito *Heamgogus spp* (Vasconcielos et al, 2003), the second cycle is the urban cycle where humans are the primary host and are transmitted by the peri urban mosquito *Aedes egypti* (Goud et al, 2003).

Yellow fever virus has one serotype, but phylogenetic analysis has identified up to 5 genotypes West Africa genotype I, East Africa genotype, East Africa/Angolan genotype, South American genotype I and South American genotype II (Chang et al, 1995). Previous genetic analysis has shown that these genotypes are geographically defined (Mutebi et al, 2001), as have been reported for other viruses (Siddle et al, 2018; Carrol et al, 2015 Nature; Motayo et al, 2019). Historical phylogenetic data show that the South American genotypes emerged from the West African genotype (Braynt et al, 2007), and have been hypothesized that this emergence was as a result of importations of infections during the slave trade (Braynt et al, 2007). Previous reports have also shown the likelihood of YF virus originating from around East Africa before emerging in West Africa. Yellow fever outbreaks continue to occur in Africa despite mass vaccination campaigns ongoing in various African countries (WHO, 2020). For instance, between 2017 to 2022 there have been outbreaks of YF in Edo, Kwara, Ebonyi and Bauchi states in Nigeria, with a total of 287 confirmed cases (Nomhwange et al, 2021; William et al, 2022). Recent outbreaks have also been reported in Senegal in 2020 with 542 suspected cases and 15 laboratory confirmed cases (Dieng et al, 2022). There have also been silent outbreaks in Nigeria, with a report among patients presenting with pyrexia of unknown origin (Ajobgasile et al, 2020). An ongoing vectorial study for arboviruses also recently identified YF virus among mosquitoes in Ibadan, Nigeria (Personal communication 2023).

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

The continuous reports of outbreaks of YF virus in different parts of Africa, despite implementation of mass vaccination campaigns calls for attention and motivated the conceptualization of this study. The current study was designed to investigate the relationship between YF field surveillance data and available genomic data to the continuous occurrence of outbreaks in Africa; as well as most recent geographic transmission events of Yellow Fever virus genotypes using updated sequence data.

METHODS

Data curation

Data relating to all reported outbreaks of yellow fever virus infections were collected after database search. The databases that were searched include PubMed, Scopus and Google Scholar. Supplementary Table 1 shows the list of Outbreaks with the corresponding articles they were reported in.

Partial YF, E gene sequences from Africa were downloaded from GenBank, with the help of the NCBI Virus database. Another South American genome sequences and 2 isolated from USA were also downloaded from GenBank. Two different datasets were generated from these sequences, the first dataset consisted genome sequences from Africa, along with reference sequence from the Americas (n = 220). The second dataset consisted of genome sequences from Africa from 1948 to 2021 (n =76), the second data set represents all archived E gene sequences from Africa, the E gene was selected because it represents the most frequently and consistently sequenced gene region for YF virus. In addition to the sequence data, metadata was also extracted such as sample source, year of isolation, and Country Location of isolation.

Outbreak Demography

The number of reported cases and subsequent deaths of infected individuals according to Country and year, were retrieved from the published reports of YF outbreaks in Africa. This data was tabulated and visualized on a global map using R studio (https://www.r-project.org). A bar and corresponding line chat representing African YF cases and deaths according to year of occurrence was also generated in R studio (https://www.r-project.org).

Phylogenetic analysis

Partial genome sequences downloaded from GenBank were aligned using MAFFTv7.222 (FF-NS-2 algorithm) following default settings (Katoh et al, 2019). Maximum likelihood phylogenetic analysis was performed using the general time reversible nucleotide substitution model with gamma distributed rate variation GTR-Γ(Yang et al, 1994) with 1000 bootstrap replicates using IQ-TREE software (Nguyen et al 2015).

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Phylogeography and Evolutionary analysis

Evolutionary and additional phylogenetic analysis was carried out on African YF strains, using a Bayesian evolutionary approach using Markov Chain Monte Carlo (MCMC) implemented in BEAST version 2.5 (Bouckaert et al, 2019). The strength of the temporal clock signal of the dataset was carried out by conducting a regression of root to tip genetic distances against year of sampling analysis using TempEst v 1.5 (Rambault et al, 2020). For the MCMC run, a total of 103 YF virus partial E gene sequences of 280 nucleotides in length were aligned, consisting 86 African sequences and 3 reference sequences from Peru and USA. A list of the sequences used for the analysis is contained in supplementary table 1. Further analysis was then done using the relaxed clock with Gausian Markov Random Field Skyride plot coalescent prior. The MCMC run was set at 100,000,000 states with a 10% burn in. Results were visualized in Tracer version 1.8. (http://tree.bio.ed.ac.uk/software/tracer/). Bayesian skyride analysis was carried out to visualize the epidemic evolutionary history using Tracer v 1.8. (http://tree.bio.ed.ac.uk/software/tracer/). Geographical coordinates of the locations of the African YF sequences were retrieved from the web with the help of online servers. A phylogeographic tree with discrete traits was constructed using the African YF sequences and their geographic coordinates in latitude and longitude using a Bayesian stochastic search variable selection (BSSVS) model implemented in BEAST 2.5 (Bouckaert et al, 2019). The resulting tree was annotated in TreeAnnotator after discarding a 10% burn-in. In addition, a continuous phylogeographic tree was also constructed for major African YF genotypes, according to their regions of circulation, WA-I/II, and EA-1/EA-II-Angolan genotypes, using a similar approach for the discrete tree except that a continuous tree trait was selected in creating the BEAST XML file in Beauti. The resulting MCC trees were visualized in ggtree (Yu et al, 2007).

Geographic dispersal

The transmission histories and geographic dispersal was projected using the Phylogeographic trees (Continuous and Discrete) were loaded unto an online software sever where information such as internal node ages, user determined geographic coordinates e.t.c. and maps these to a world geographic map. rendered into SPREAD4 (Nahata et al, 2022).

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

RESULTS

Yellow fever outbreak Demography

African Yellow fever demography was re-evaluated with updated data from recent literature (WHO 2016; Nomhwange et al, 2021). There were also silent outbreaks identified from molecular detection from patients with pyrexia of unknown origin in Nigeria (Ajobgasile et al, 2020). From our analysis, Nigeria remains the country with highest number of cases, with Ethiopia having the least number of cases (Figure 1a). The case fatality distribution shows that high mortality rates were recorded between 1984 and 1987 Figure 1b, with majority of cases arising from Nigeria and Ghana Supplementary table 1.

Phylogeography and Evolution

Global phylogeny of YF virus revealed that the African viruses clustered within the four major genotypes West African genotype I (WA-I), West African genotype II (WA-II), East African genotype (EA), Angolan genotype or East African genotype II (EA-II). Majority of the African strains fell under the WA-II, while WA-I had the lowest number of viral sequences as shown in Figure 3. There was strict geographic defined phylogenetic diversification as shown by continental clustering of the viruses analyzed in our study with the South American strains clustering within the South American genotype and African strains within the African genotypes.

The root to tip divergence for the African YF sequences showed a positive signal, with R^2= 0.11. The MCC tree of the African YF isolates showed a similar diversification with the global tree with Senegalese isolates being the most abundant and diversly distributed within both WA-I and WA-II genotype sub-clades (Figure 5). The evolutionary rate of the African YF viruses was 2.08 X 10^{-4}, 98% HPD (8.48 X 10^{-5} - 3.41 X10^{-3}) substitutions/ site/ year. The time calibrated MCC tree also showed ancestral divergence times for each ancestral node with the earliest ancestor dating back as the year 1140. The WA-I genotype had a time to most recent ancestor TMRCA of 1838, 95% HPD (1740 – 1910), while WA-II had a TMRCA of 1840, 95% HPD (1730 – 1920). The EA genotype on the other hand had a TMRCA of 1720, 95% HPD (1740 – 1910), and EA-II 1940, 95% HPD (1900 – 1970) as shown in Supplementary Figure 1. Other ancestral nodes for the genotypes, particularly WA genotypes displayed high posterior support (supplementary Figure 1).

Evolutionary rate of the African YF sequences was. The viral population of the African sequences had a steady growth between 1930 to 1970, and a sharp decline in effective population between 1970 to the most recent sampling date.

Yellow fever virus geographic dispersal

We explored the pattern of geographic dispersal and viral circulation of YF virus within Africa using both discrete and continuous phylogeographic reconstructions. The discrete model clearly

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

showed major transmission routes within West Africa particularly, areas covering countries such as Nigeria and Senegal. Transmission pathways were also seen within East and Central Africa. The phylogeographic movement seemed to be localized within the two regions with the West African transition events moving westward, while the East African transitions were mainly eastward movements. The continuous phylogeography showed a series of recent transmission events represented by the lighter shaded polygons (Figure 5) concentrated within countries such as North Eastern Nigeria to Cameroon and Central Africa Republic, and between Northern Nigeria across parts of Niger republic up to Burkina Fasso in West Africa. There were also recent dispersal events observed around Uganda, Kenya and parts of Tanzania and even into the horn of Africa such as Ethiopia.

DISSCUSION

General and Molecular Epidemiology

Yellow fever virus was first reported in America around the 15th century by Spanish conquerors, but historical and molecular evidence indicates that the virus first originated in Africa where it was imported into the Americas through the trans-Atlantic slave trade (White, 1959; Braynt et al, 2007), ever since then there has been several large outbreaks in mostly West Africa and parts of Central Africa (Chippaux et al, 2018). Yellow fever virus was first isolated from the blood of a Ghanian man called Asibi in 1927 (Stapes and Monath, 2008). Epidemiological data from previous outbreaks, between 1979 and 2018 reviewed in our study indicates a fluctuating pattern of intermittent outbreaks with its peak at between 1986 and 1988 (Figure 1b). There was also an upsurge in cases resulting from a large outbreak in Angola in 2015, followed by Uganda 2016. It is presumed that certain social drivers were largely responsible for this re-emergence of large epidemics in East Africa, these include long term conflict among waring rebels and government forces, human population displacement resulting from conflicts, and improved surveillance activity.

Molecular epidemiology of the YF virus was consistent with previous reports (Mutebi et al, 2001; Sall et al, 2010; Grobbelaar et al, 2016). With the African sequences maintaining their geographically defined genotypes (Figure 2). There was one West African genotype II reported from the Central African Republic GenBank accession number GU073130. The virus was isolated from a mosquito in 1974 in Central African Republic, an indication that the WA-II genotype might have been circulating and maintained in the forests of CAR. The EA genotypes I and II were strictly maintained in their geographical region. The majority of the recent sequences were of WA-II, with the genotype dominating the Entire West Africa and parts of Central Africa.

Phylogeography and transmission dynamics

Yellow fever phylogeography have consistently showed a distinct geographically bound diversification (Mutebi et al, 2001; Sall et al, 2010; Li 2022). Our present study maintains the

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

same regional clustering of the West African genotypes and the East/Central African genotype as well as the Angolan genotype. The more recent sequences isolated from Nigeria in 2018, clustered within WA-II, having a common ancestral history as other closely related isolates (Supplementary Fig 2), sharing the same sub-clade with isolates from Senegal, Cameroon, and Burkina Fasso. This observation shows strict geographical conservation probably due to the fact that same genotype has established itself within the Forests and grasslands of Nigeria all the way to Senegal. Previous studies have hypothesized that sustained epizootic transmission cycles within forested regions of Africa have led to YF virus adaptation to its intermediate host NHP and its competent vector Heamagogous spp (Mutebi et al, 2001). Also, the fact that humans act as a dead-end host leaves little opportunity for intra-host evolution and diversity among human population. Previous reports have identified key mutational features among African YF isolates that confer a positive selection pressure and may play a role in genotype emergence (Vasconcelos et al, 2004; Brayant et al, 2007). The EA-I and the Angolan genotypes were in conformity with previously reported geographically defined diversity (Mutebi et al, 2001; Sall et al, 2010). The most recently identified strains from outbreak in Angola and Uganda 2015/2016 clustered within separate genotypes.

Discrete phylogeography showed active transmission routes within both West Africa and East Africa, as against between both regions. This is an indication of sustained intra-regional transmission as against trans regional transmission despite constant human movement between West and Central Africa. This observation serves as evidence that YF transmission dynamics is sparsely influenced by human activity such as international trade or human migration, but is more likely affected by ecological drivers and climatic events that may affect the Forest /Savanah ecosystem as previously reported (Moriera-Soto et al, 2018; Hill et al, 2020). Continuous phylogeography revealed a robust temporal spread among the genotypes within their regions of primary emergence. The WA genotypes seemed to have stronger migration events compared to the EA genotypes. The WA genotypes displayed a westward movement cutting across more than 5 countries all the way from Central Africa Republic to Cameroon across Nigeria to as far west as Senegal. This affirms that vegetation plays a vital role in the geographic spread of YF across Africa as there was limited Northward geographic movement of the WA genotypes. The EA/Angolan genotypes seemed to maintain their dispersal activity within the East/Central African region, although some northward movement was observed with spread events into Sudan and Ethiopia, these were just probably isolated occurrences arising from local outbreaks from Sylvatic/Sahvana YF cycle spillovers into human habitation areas.

CONCLUSION

This study reports the most current data on YF virus phylogeography and transmission dynamics. We have shown that YF virus maintain a strict geographically bound lineage dispersal pattern, with major genotypes circulating within strict regional confines. We also report a major surge in reported laboratory confirmed cases majorly in East Africa, after the year 2015. There was also rapid expansion of the WA-II genotype in the last 4 years. Study limitations include poor

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

genomic surveillance system across Africa, resulting in limited amount of YF sequence data available in open-source data banks. We therefore recommend expanded human and vectoral surveillance of YF outbreaks as well as other Arboviruses of public health importance. We also advocate for upscaling Molecular biology capabilities of existing public health labs both Nationally and regionally to enable rapid genomic sequencing and analysis of samples from Arboviral outbreaks such as YF to help in prevention of future epidemics and protect Africa from existing biological threats.

FIGURE LEGENDS

Figure 1a Geographic distribution of YF virus cases across Africa 1979 to 2019, color intensity is representative of the number of confirmed YF cases in each country. **Figure 1b** Table showing the number of confirmed Yellow Fever cases versus the number of deaths (Red line) from 1979 to 2019.

Figure 2. Maximum Likelyhood tree of African Yellow Fever isolates along with some global reference sequences.

Figure 3a. Root to tip regression analysis of African Yellow Fever sequences. **Figure 3b**. Time resolved Phylogeographic MCC tree of African Yellow virus E gene sequences.

Figure 4. Spatiotemporal reconstruction of the discrete phylogeography of Yellow Fever virus in Africa. Curved lines depict links between the nodes represented by circles. The pink circles are the nodes representing the maximum likelihood phylogeny. Directionality of the movements is anticlockwise along the curves.

Figure 5a. Continuous phylogeography of West African Yellow Fever genotypes, shaded blue circles represent 80% highest posterior density interval of the respective nodes of the MCC tree. The lighter shaded area represents more recent viral introductions. Directionality of the movements is anticlockwise along the curves. **Figure 5b**. Continuous phylogeography of East African/Angolan Yellow Fever genotypes, shaded blue circles represent 80% highest posterior density interval of the respective nodes of the MCC tree. The lighter shaded area represents more recent viral introductions. Directionality of the movements is anticlockwise along the curves.

FINANCIAL DISCLOSURE

The authors did not receive any funding or financial support for this project.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of the paper.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

REFERENCES

Ajogbasile, F. V., Oguzie, J. U., Oluniyi, P. E., Eromon, P. E., Uwanibe, J. N., Mehta, S. B., Siddle, K. J., Odia, I., Winnicki, S. M., Akpede, N., Akpede, G., Okogbenin, S., Ogbaini-Emovon, E., MacInnis, B. L., Folarin, O. A., Modjarrad, K., Schaffner, S. F., Tomori, O., Ihekweazu, C., Sabeti, P. C., … Happi, C. T. (2020). Real-time Metagenomic Analysis of Undiagnosed Fever Cases Unveils a Yellow Fever Outbreak in Edo State, Nigeria. *Scientific reports*, *10*(1), 3180. https://doi.org/10.1038/s41598-020-59880-w.

Beasley D.W., McAuley A.J., Bente D.A. Yellow fever virus: Genetic and phenotypic diversity and implications for detection, prevention and therapy. Antivir. Res. 2015;115:48–70. doi: 10.1016/j.antiviral.2014.12.010.

Bryant JE, Holmes EC, Barrett AD. Out of Africa: a molecular perspective on the introduction of yellow fever virus into the Americas. PLoS Pathog. 2007 May 18;3(5):e75. doi: 10.1371/journal.ppat.0030075.

Bouckaert R., Vaughan T.G., Barido-Sottani J., Duchêne S., et al. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS computational biology, 15(4), e1006650.

Carroll, M. W., Matthews, D. A., Hiscox, J. A., Elmore, M. J., Pollakis, G., Rambaut, A., Hewson, R., García-Dorival, I., Bore, J. A., Koundouno, R., Abdellati, S., Afrough, B., Aiyepada, J., Akhilomen, P., Asogun, D., Atkinson, B., Badusche, M., Bah, A., Bate, S., Baumann, J., … Günther, S. (2015). Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. *Nature*, *524*(7563), 97–101.

Chippaux J.P., Chippaux A. Yellow fever in Africa and the Americas: A historical and epidemiological perspective. J. Venom. Anim. Toxins. Incl. Trop. Dis. 2018;24:20. doi: 10.1186/s40409-018-0162-y.

Chang, G. J., Cropp, B. C., Kinney, R. M., Trent, D. W., & Gubler, D. J. (1995). Nucleotide sequence variation of the envelope protein gene identifies two distinct genotypes of yellow fever virus. *Journal of virology*, *69*(9), 5773–5780.

Dieng, I., Diallo, A., Ndiaye, M., Mhamadi, M., Diagne, M. M., Sankhe, S., Ndione, M. H. D., Gaye, A., Sagne, S. N., Heraud, J. M., Sall, A. A., Fall, G., Loucoubar, C., Faye, O., & Faye, O. (2022). Full genome analysis of circulating DENV-2 in Senegal reveals a regional diversification into separate clades. *Journal of medical virology*, *94*(11), 5593–5600.

Gould EA, de Lamballerie X, Zanotto PM, Holmes EC. 2003. Origins, evolution, and vector/host coadaptations within the genus Flavivirus. Adv. Virus Res. 59:277–314.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Grobbelaar, A. A., Weyer, J., Moolla, N., Jansen van Vuren, P., Moises, F., & Paweska, J. T. (2016). Resurgence of Yellow Fever in Angola, 2015-2016. *Emerging infectious diseases*, *22*(10), 1854–1855.

Hasegawa M, Kishino H, Yano T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 22:160–74. doi:10.1007/BF02101694.

Hill, S. C., de Souza, R., Thézé, J., Claro, I., Aguiar, R. S., Abade, L., Santos, F. C. P., Cunha, M. S., Nogueira, J. S., Salles, F. C. S., Rocco, I. M., Maeda, A. Y., Vasami, F. G. S., du Plessis, L., Silveira, P. P., de Jesus, J. G., Quick, J., Fernandes, N. C. C. A., Guerra, J. M., Réssio, R. A., … Faria, N. R. (2020). Genomic Surveillance of Yellow Fever Virus Epizootic in São Paulo, Brazil, 2016 - 2018. *PLoS pathogens*, *16*(8), e1008699. https://doi.org/10.1371/journal.ppat.1008699

Katoh K, Misawa K, Kuma K, Miyata T , MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, *Nucleic Acids Research*, Volume 30, Issue 14, 15 July 2002, Pages 3059–3066.

Li Y. (2022). Molecular epidemiology of yellow fever virus in Africa: A perspective of the phylogeographic split between East/Central African and West African lineages. *Acta tropica*, *225*, 106199. https://doi.org/10.1016/j.actatropica.2021.106199

Monath TP, Vasconcelos PFC. Yellow fever. Journal of Clinical Virology. 2015;64: 160–173. 10.1016/j.jcv.2014.08.030.

Motayo, B. O., Oluwasemowo, O. O., Olusola, B. A., Opayele, A. V., & Faneye, A. O. (2019). Phylogeography and evolutionary analysis of African Rotavirus a genotype G12 reveals district genetic diversification within lineage III. *Heliyon*, *5*(10), e02680.

Moreira-Soto A, Torres MC, Mendonc¸a MCL de, Mares-Guia MA, Rodrigues CD dos S, Fabri AA, et al. Evidence for multiple sylvatic transmission cycles during the 2016–2017 yellow fever virus outbreak, Brazil. Clinical Microbiology and Infection. 2018; 24: 1019.e1–1019.e4.

Mutebi, J. P., Wang, H., Li, L., Bryant, J. E., & Barrett, A. D. (2001). Phylogenetic and evolutionary relationships among yellow fever virus isolates in Africa. *Journal of virology*, *75*(15), 6999–7008.

Nahata, K. D., Bielejec, F., Monetta, J., Dellicour, S., Rambaut, A., Suchard, M. A., Baele, G., & Lemey, P. (2022). SPREAD 4: online visualisation of pathogen phylogeographic reconstructions. *Virus evolution*, *8*(2), veac088. https://doi.org/10.1093/ve/veac088.

Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* *32*, 268–274. https://doi.org/10.1093/molbev/msu300.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Nomhwange, T., Jean Baptiste, A. E., Ezebilo, O., Oteri, J., Olajide, L., Emelife, K., Hassan, S., Nomhwange, E. R., Adejoh, K., Ireye, F., Nora, E. E., Ningi, A., Bathondeli, B., & Tomori, O. (2021). The resurgence of yellow fever outbreaks in Nigeria: a 2-year review 2017-2019. *BMC infectious diseases*, *21*(1), 1054. https://doi.org/10.1186/s12879-021-06727-y.

Nwachukwu William E, Oladejo J, Ofoegbunam CM, Anueyiagu C, Dogunro F, Etiki SO, Dachung BI, Obiekea C, Aderoju B, Akanbi K, Adeyemi IT, Famokun GA, Emelife O, Osamwonyi IW, Ochu CL, Abiode A, Ireye F, Chukwuji M, Ipadeola O, Saiki M, Okudo I, Nwodo D, Oteri JA, Ilori E, Mba N, Ihekweazu C. Epidemiological description of and response to a large yellow fever outbreak in Edo state Nigeria, September 2018 - January 2019. BMC Public Health. 2022 Aug 30;22(1):1644. doi: 10.1186/s12889-022-14043-6.

Rambaut, A. Lam, T.T. Carvalho, L.M. Pybus, O.G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2020. 2, vew007.

Siddle, K. J., Eromon, P., Barnes, K. G., Mehta, S., Oguzie, J. U., Odia, I., Schaffner, S. F., Winnicki, S. M., Shah, R. R., Qu, J., Wohl, S., Brehio, P., Iruolagbe, C., Aiyepada, J., Uyigue, E., Akhilomen, P., Okonofua, G., Ye, S., Kayode, T., Ajogbasile, F., … Happi, C. T. (2018). Genomic Analysis of Lassa Virus during an Increase in Cases in Nigeria in 2018. *The New England journal of medicine*, *379*(18), 1745–1753.

Staples, J. E., & Monath, T. P. (2008). Yellow fever: 100 years of discovery. *JAMA*, *300*(8), 960–962. https://doi.org/10.1001/jama.300.8.960.

Vasconcelos PFC, Sperb AF, Monteiro HAO, Torres MAN, Sousa MRS, Vasconcelos HB, et al. Isolations of yellow fever virus from Haemagogus leucocelaenus in Rio Grande do Sul State, Brazil. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2003; 97: 60–62.

Vasconcelos PF, Bryant JE, da Rosa TP, Tesh RB, Rodrigues SG, et al. (2004) Genetic divergence and dispersal of yellow fever virus, Brazil. Emerg Infect Dis. 10: 1578–1

White C.R. Yellow fever; history of the disease in the eighteenth and nineteenth century. J. Kans. Med. Soc. 1959;60:298–302.

WHO Countries with risk of yellow fever transmission and countries requiring yellow fever vaccination. https://www.who.int/publications/m/item/countries-with-risk-of-yellowfever-transmission-and-countries-requiring-yellow-fever-vaccination-(july-2020).

Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T.T.-Y. ggtree: An r package for visualization and annotation of phylogenetic treeswith their covariates and other associated data. Methods Ecol. Evol. 8, 28–36 (2017).

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

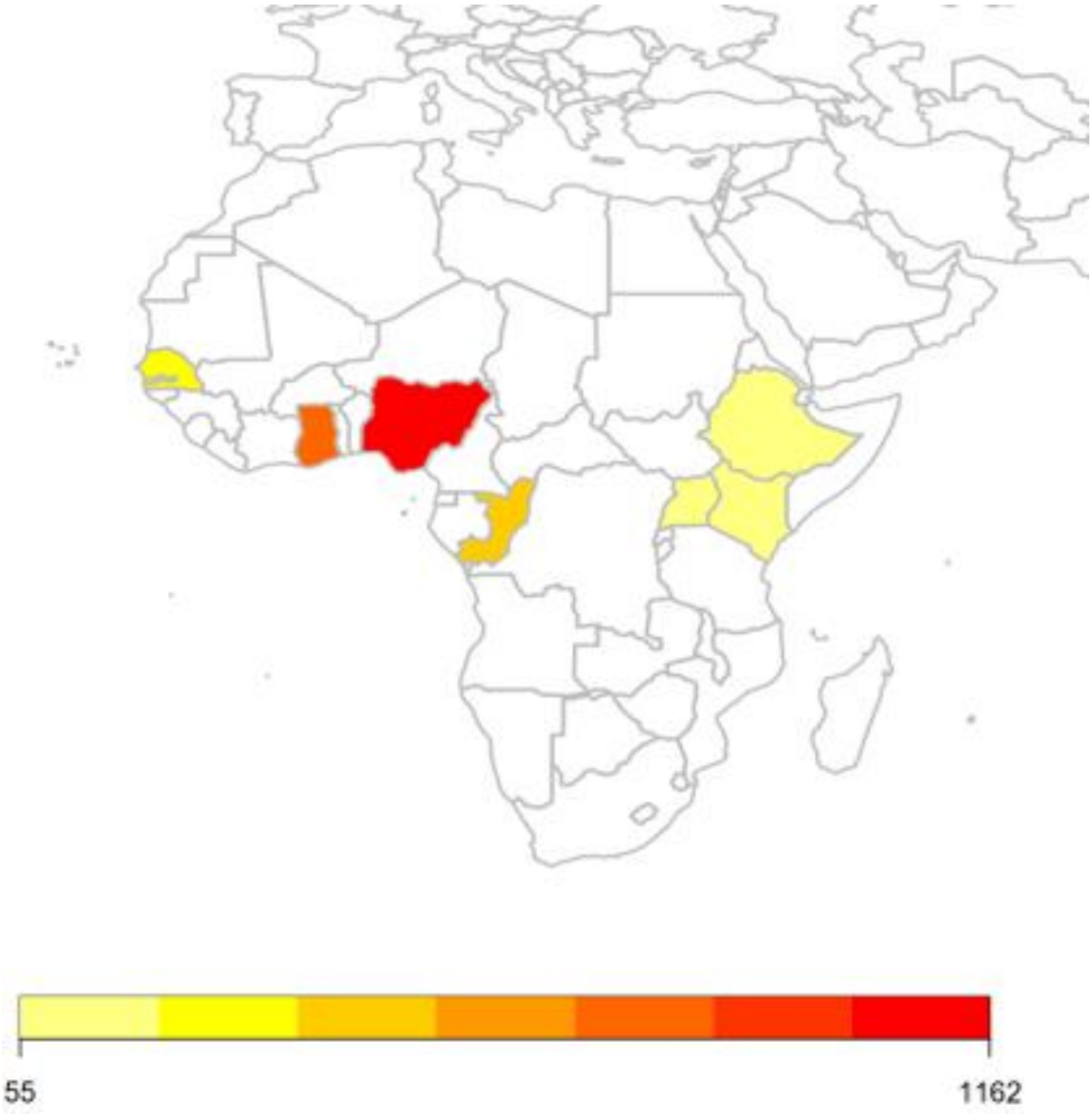

Figure 1a Geographic distribution of YF virus cases across Africa 1979 to 2019, color intensity is representative of the number of confirmed YF cases in each country.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

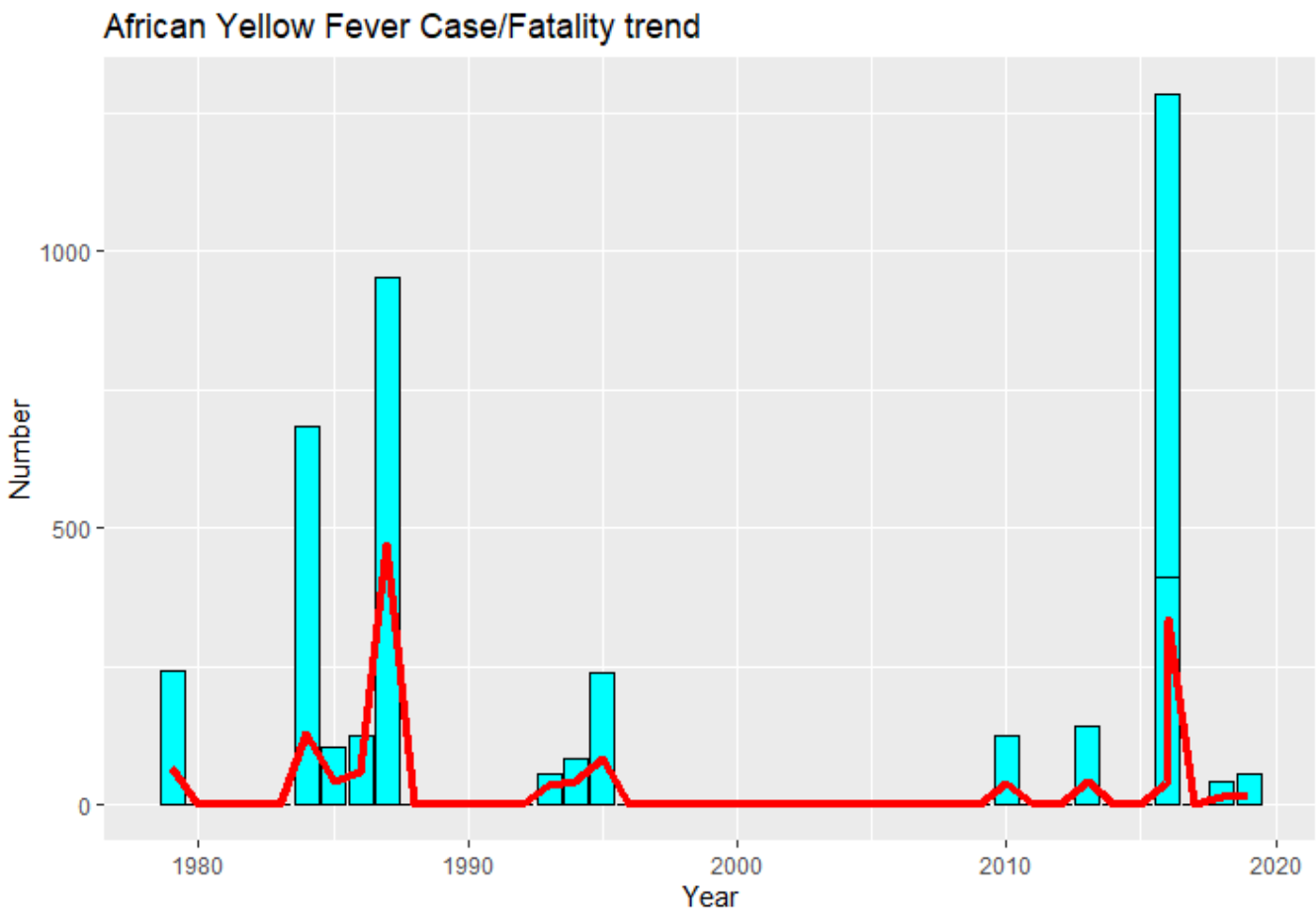

Figure 1b Table showing the number of confirmed Yellow Fever cases versus the number of deaths (Red line) from 1979 to 2019.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

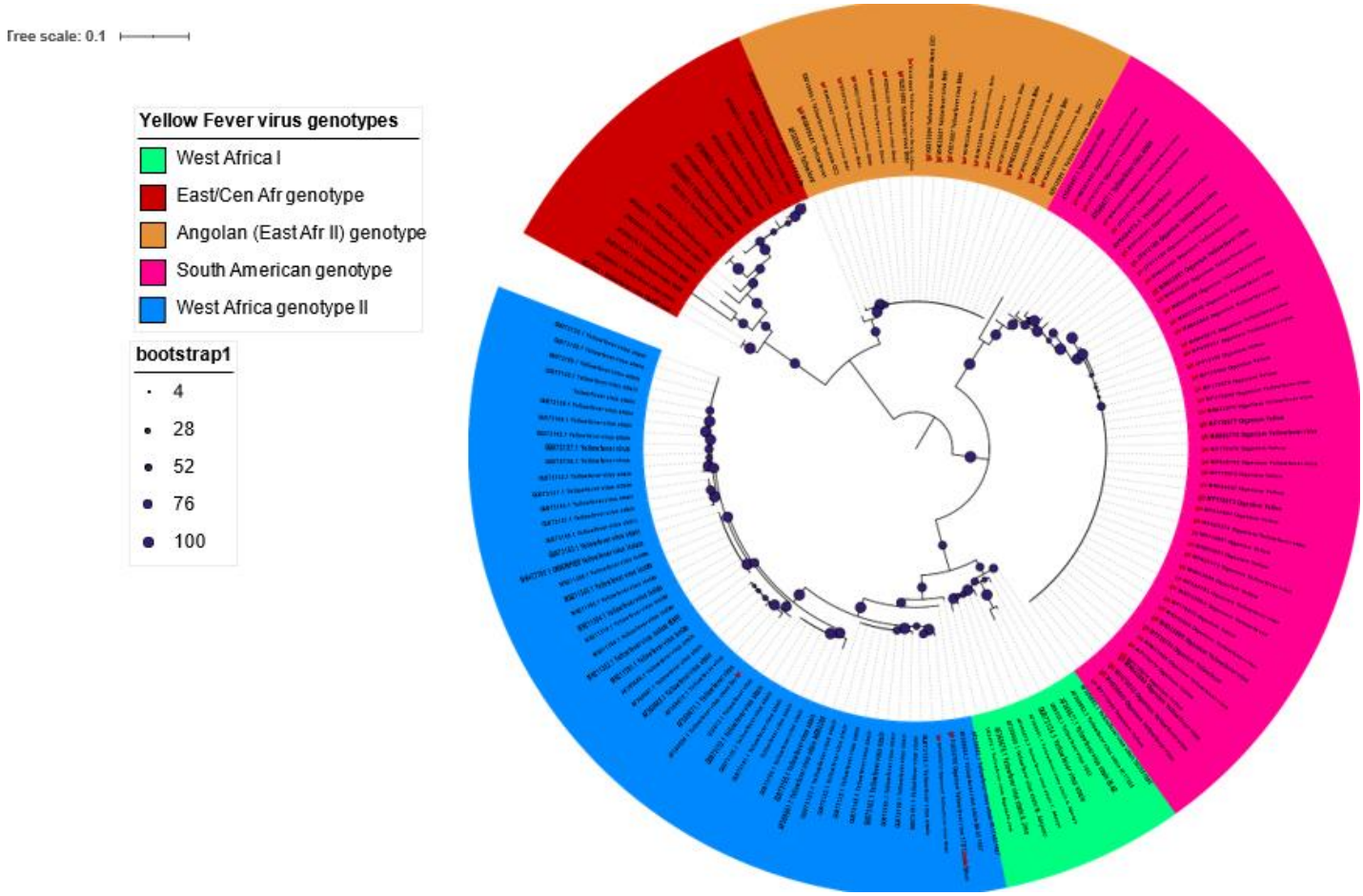

Figure 2. Maximum Likelyhood tree of African Yellow Fever isolates along with some global reference sequences.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

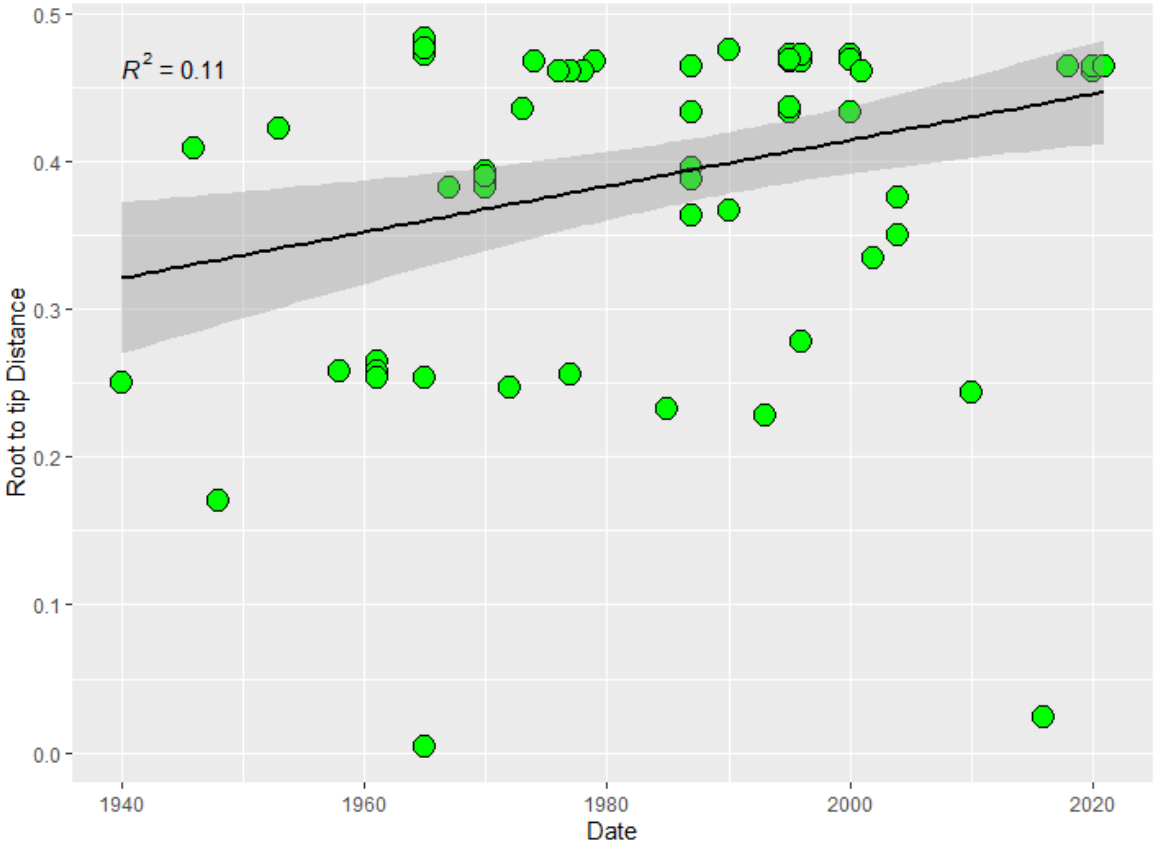

Figure 3a. Root to tip regression analysis of African Yellow Fever sequences

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

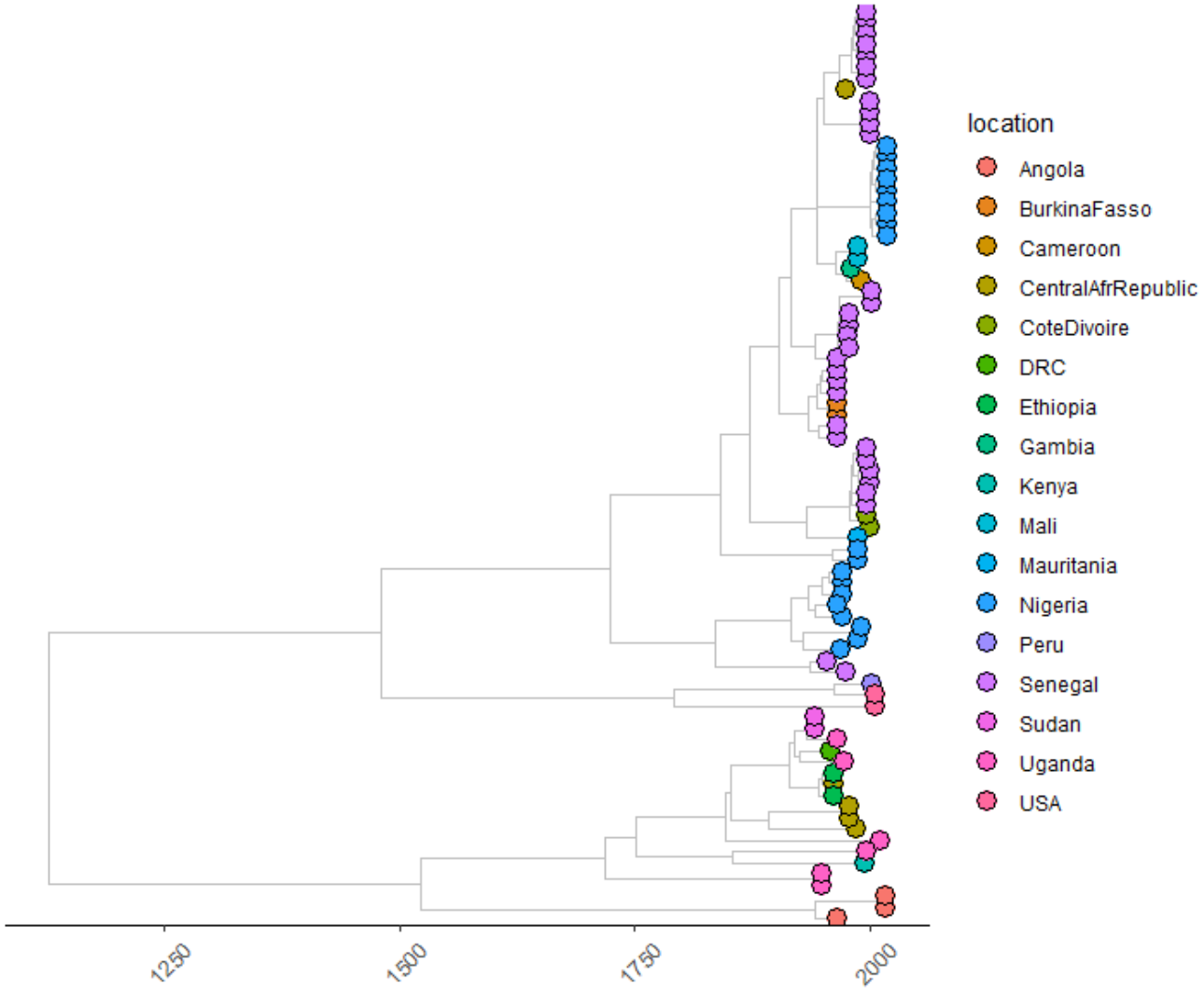

Figure 3b. Time resolved Phylogeographic MCC tree of African Yellow virus E gene sequences

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Figure 4. Spatiotemporal reconstruction of the discrete phylogeography of Yellow Fever virus in Africa. Curved lines depict links between the nodes represented by circles. The pink circles are the nodes representing the maximum likelihood phylogeny. Directionality of the movements is anticlockwise along the curves.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

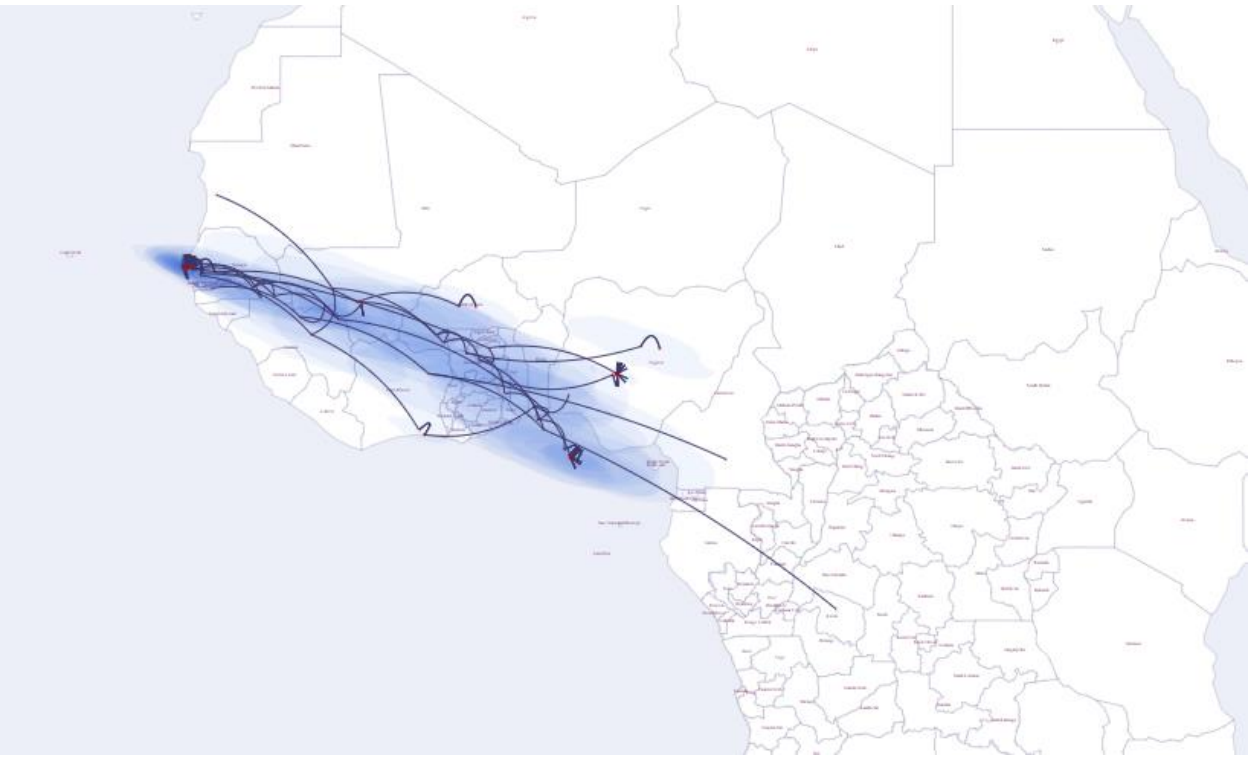

Figure 5a. Continuous phylogeography of West African Yellow Fever genotypes, shaded blue circles represent 80% highest posterior density interval of the respective nodes of the MCC tree. The lighter shaded area represents more recent viral introductions. Directionality of the movements is anticlockwise along the curves.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.06.23291042; this version posted June 7, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

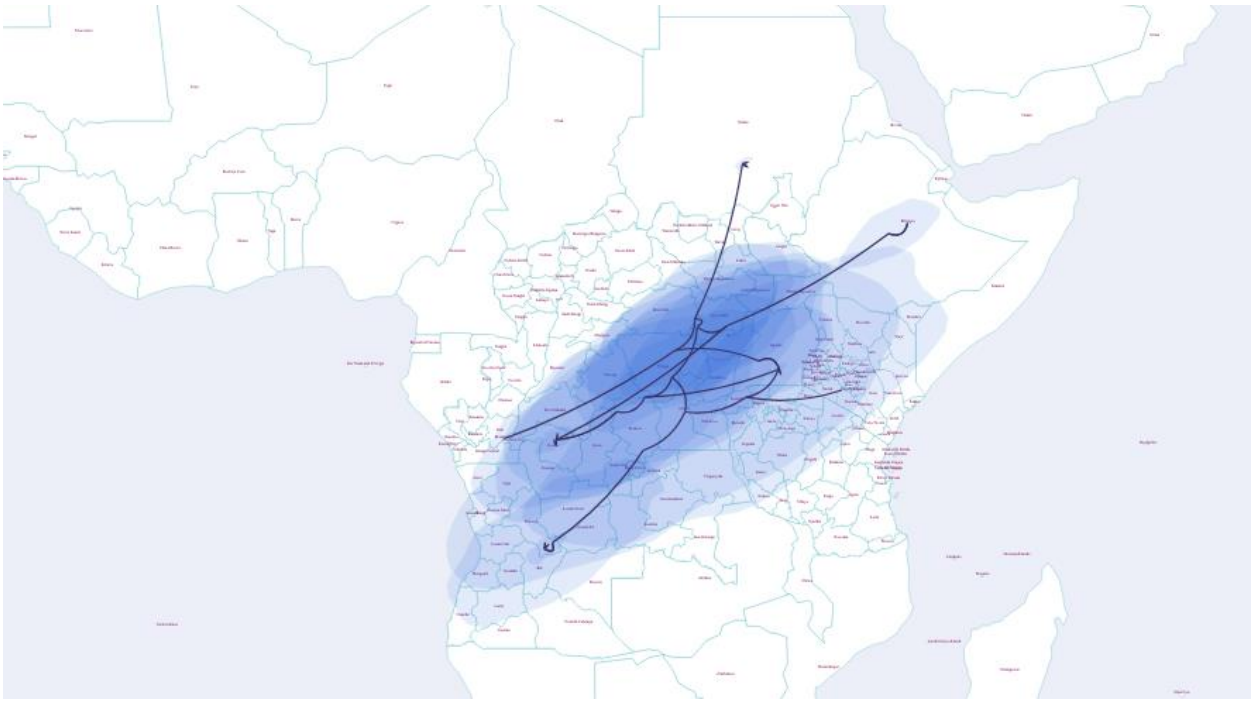

Figure 5b. Continuous phylogeography of East African/Angolan Yellow Fever genotypes, shaded blue circles represent 80% highest posterior density interval of the respective nodes of the MCC tree. The lighter shaded area represents more recent viral introductions. Directionality of the movements is anticlockwise along the curves.