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], Isibor Patrick Omoregie[3], Solomon Uche Oranusi[3]

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 ; Paul A Akinduti niyiakinduti@gmail.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 paucity of data still exists regarding directional spread and phylogeography of the African Yellow fever genotypes. The need to fill this gap with information from spatiotemporal data 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 in Africa, to analyze the case/fatality distribution. 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 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.

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 [1]. 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 [2]. Yellow fever is an arthropod borne infection, maintained by the mosquito vector *Aedes aegypti*, and *Aedes albopticus*, although several vectorial studies have identified YF virus in other mosquitoes' species, such as Aedes *taylori* [3]. 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 [3]. 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* [4], the second cycle is the urban

cycle where humans are the primary host and are transmitted by the peri urban mosquito *Aedes aegypti* [3].

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 [5]. Previous genetic analysis has shown that these genotypes are geographically defined [6], as have been reported for other viruses [7-9]. Historical phylogenetic data show that the South American genotypes emerged from the West African genotype [10], and have been hypothesized that this emergence was as a result of importations of infections during the slave trade [10]. 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 [11]. 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 [12,13]. Recent outbreaks have also been reported in Senegal in 2020 with 542 suspected cases and 15 laboratory confirmed cases [14]. There have also been silent outbreaks in Nigeria, with a report among patients presenting with pyrexia of unknown origin [15]. An ongoing vectorial study for arboviruses also recently identified YF virus among mosquitoes in Ibadan, Nigeria (Personal communication 2023).

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, along with sequences that were recovered from returnee traveler from Angola to China after the 2016 Angolan YF outbreak that spread to China via international travel [16]. This was done with the help of the NCBI Virus database. The dataset consisted of genome sequences of all archived E gene sequences from Africa, along with whole genome sequences deposited in GenBank, 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.

Demographic analysis

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 representing African YF cases and deaths according to year of occurrence and Country 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 [17]. The dataset consisted genome sequences from Africa available in GenBank as at 1st February 2023, along with representative sequences from the Southern America (n =220). Maximum likelihood phylogenetic analysis was performed using the general time reversible nucleotide substitution model with gamma distributed rate variation GTR-Γ[18] with 1000 bootstrap replicates using IQ-TREE software [19].

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 [20]. 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 [21]. For the MCMC run, a total of 110 YF virus partial E gene sequences of 280 nucleotides in length were aligned, consisting 91 African sequences and 19 Chinese sequences from the international outbreak in 2016. 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 [20]. 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 [22].

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 [23].

RESULTS

Yellow fever outbreak Demography

African Yellow fever demography was re-evaluated with updated data from recent literature [12, 24]. There were also silent outbreaks identified from molecular detection from patients with pyrexia of unknown origin in Nigeria [15]. From our analysis, Nigeria remains the country with highest number of cases, with Ethiopia having the least number of cases (Figure 1top). The case fatality distribution shows that high mortality rates were recorded between 1984 and 1987 Figure 1 bottom, with majority of cases arising from Nigeria, Angola 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 2. 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= 0.079 (Figure 3 top). 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 3 bottom). 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 1845, 95% HPD (1795 – 1890), while WA-II had a TMRCA of

1852, 95% HPD (1790 – 1910). The EA genotype on the other hand had a TMRCA of 1760, 95% HPD (1655 – 1850), and EA-II 1940, 95% HPD (1900 – 1970). Other ancestral nodes for the genotypes, particularly WA genotypes displayed high posterior support (supplementary Figure 1).

The viral population of the African sequences had a steady declined between 1930 to 1970, with a sharp decline in effective population between 1970 to the most recent sampling date (Supplementary Figure 2).

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 showed major transmission routes between East and West Africa. There was also Transmission pathways were also seen within East and Central Africa at 0.5 Bayes Factor (BF). At a lower BF there were molecular jumps between East Africa and China (Figure), representing transmission events but were not observed at higher BF (Figure). The phylogeographic movement seemed to be localized between the two regions of West and East African, with multiple transition events moving westward. The continuous phylogeography showed a series of recent viral introductions from historically endemic areas; represented by the lighter shaded polygons (Figure 5) such as North Eastern Nigeria and parts of South Eastern Nigeria and Cameroon in West Africa. Recent viral introductions were observed in Ethiopia in East Africa, although there were recent introductions in countries such as Angola, but the polygons seemed to overlap around these regions owing to multiple outbreaks over time and the paucity of the data. The color gradient goes from recent (dark) to ancient (light) colored lines along the transition routes of the dispersal events. From this representation, recent dispersal events observed around Democratic Republic of Congo (DRC), Kenya, and parts of Tanzania and even into the horn of Africa such as Ethiopia, all in East and Central Africa.

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 [10, 25], ever since then there has been several large outbreaks in mostly West Africa and parts of Central Africa [26]. Yellow fever virus was first isolated from the blood of a Ghanian man called Asibi in 1927 [27]. Epidemiological data from previous outbreaks, between 1960 and 2022 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, which spilled over into the Democratic Republic of Congo, 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 [6, 28, 29]. 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

There is a paucity of data regarding spacial epidemiology of Infectious agents in Africa, recent reports from West Africa have shown the significance of molecular biogeography and special epidemiology in tracking rapidly expanding infectious agents within West Africa [30-32]. Yellow fever phylogeography have consistently showed a distinct geographically bound diversification [6, 28, 33]. Our present study maintains the 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 [6]. 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 [10, 34]. The EA-I and the EA-II/Angolan genotypes were in conformity with previously reported geographically defined diversity [6, 28]. 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 and between both regions. This is an indication of sustained intra-regional transmission as well as trans regional transmission resulting from constant human movement between West and Central Africa. This observation serves as evidence that YF transmission dynamics is not just

influenced by ecological drivers and climatic events that may affect the Forest /Savanah ecosystem as previously reported [35, 36], but also by human activity such as international trade and human migration. 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 as depicted by numerous recent transmission events towards the Senegalese region (darker colored lines). 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 (Figure 5a). 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 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 1 (top) Geographic distribution of YF virus cases, and deaths respectively across Africa, color intensity is representative of the number of confirmed YF cases in each country and is displayed as a color gradient by the side of each map. The NA represents not available.
1 (bottom) Bar charts showing the number of confirmed Yellow Fever cases and deaths from

1970 to 2021 according to Country. The legend is color coded to represent the various countries where outbreaks occurred.

Figure 2. Maximum Likelyhood tree of African Yellow Fever isolates along with some global reference sequences. The legend represents the various global YF genotypes.

Figure 3 (top). Root to tip regression analysis of African Yellow Fever sequences. **3 (bottom)**. Time resolved Phylogeographic MCC tree of African Yellow virus E gene sequences, showing the various Countries of origin as tips. The genotypes are represented by horizontal bars by the side of the tree. WA represents West Africa, EA East Africa, and A Angola.

Figure 4. Spatiotemporal reconstruction of the discrete phylogeography of Yellow Fever virus in Africa. Curved lines depict links between the nodes. The Bayes Factor is set at 0.3 (top) and 0.5 (bottom). Directionality of the movements is anticlockwise downward curve left to rightward movement; upward curve is right to leftward movement.

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, downward curve left to rightward movement; upward curve is right to leftward movement. **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, downward curve left to rightward movement; upward curve is right to leftward movement.

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.

REFERENCES

1. Monath TP, Vasconcelos PFC. Yellow fever. Journal of Clinical Virology. 2015;64: 160–173. 10.1016/j.jcv.2014.08.030.
2. 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.

3. 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.
4. 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.
5. Chang, G. J., Cropp, B. C., Kinney, R. M., Trent, D. W., & Gubler, D. J. Nucleotide sequence variation of the envelope protein gene identifies two distinct genotypes of yellow fever virus. Journal of virology, 1995.69(9), 5773–5780.
6. Mutebi, J. P., Wang, H., Li, L., Bryant, J. E., & Barrett, A. D. Phylogenetic and evolutionary relationships among yellow fever virus isolates in Africa. Journal of virology, 2001.*75*(15), 6999–7008. doi:10.1128/JVI.75.15.6999-7008.2001.
7. 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. Genomic Analysis of Lassa Virus during an Increase in Cases in Nigeria in 2018. *The New England journal of medicine*, 2018. *379*(18), 1745–1753. doi:10.1056/NEJMoa1804498.

8. 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. Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. *Nature*,2015. *524*(7563), 97–101. doi:10.1038/nature14594.

9. Motayo, B. O., Oluwasemowo, O. O., Olusola, B. A., Opayele, A. V., & Faneye, A. O. Phylogeography and evolutionary analysis of African Rotavirus a genotype G12 reveals district genetic diversification within lineage III. *Heliyon*, 2019. *5*(10), e02680. doi:10.1016/j.heliyon.2019.e02680.

10. 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.
11. WHO Countries with risk of yellow fever transmission and countries requiring yellow fever vaccination.2020. https://www.who.int/publications/m/item/countries-with-risk-of-yellowfever-transmission-and-countries-requiring-yellow-fever-vaccination-(july-2020).
12. 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. The resurgence of yellow fever outbreaks in Nigeria: a 2-year review

2017-2019. *BMC infectious diseases*, 2021. *21*(1), 1054. https://doi.org/10.1186/s12879-021-06727-y.

13. Nwachukwu WE, Oladejo J, Ofoegbunam CM, Anueyiagu C, Dogunro F, Etiki SO, Dachung BI, Obiekea C, et al. Epidemiological description of and response to a large yellow fever outbreak in Edo state Nigeria, September 2018 - January 2019. BMC Public Health. 2022. 22(1):1644. doi: 10.1186/s12889-022-14043-6.
14. 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. Full genome analysis of circulating DENV-2 in Senegal reveals a regional diversification into separate clades. *Journal of medical virology*,2022. *94*(11), 5593–5600.
15. Ajogbasile, F. V., Oguzie, J. U., Oluniyi, P. E., Eromon, P. E., Uwanibe, J. N., Mehta, S. B., et al. Real-time Metagenomic Analysis of Undiagnosed Fever Cases Unveils a Yellow Fever Outbreak in Edo State, Nigeria. *Scientific reports*, 2020. *10*(1), 3180. https://doi.org/10.1038/s41598-020-59880-w.
16. Li C, Li D, Smart S.J, Zhou L, Yang P, Ou J, et al. Evaluating the importation of Yellow Fever cases into China in 2016 and strategies used to prevent and control the spread of the disease. 2020. WPSAR. 11(2): doi: 10.5365/wpsar.2018.9.1.007.
17. 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.
18. Yang Z. Estimating the pattern of nucleotide substitution. *Journal of molecular evolution*, 1994. *39*(1), 105–111.
19. 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.2015. 32, 268–274. https://doi.org/10.1093/molbev/msu300.
20. Bouckaert R., Vaughan T.G., Barido-Sottani J., Duchêne S., et al. . BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS computational biology, 2019. 15(4), e1006650. doi:10.1371/journal.pcbi.1006650.
21. 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.
22. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T.T.-Y. ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 2017. 8, 28–36.
23. Nahata, K. D., Bielejec, F., Monetta, J., Dellicour, S., Rambaut, A., Suchard, M. A., Baele, G., & Lemey, P. SPREAD 4: online visualisation of pathogen phylogeographic reconstructions. Virus evolution, 2022. *8*(2), veac088. https://doi.org/10.1093/ve/veac088.

24. World Health Organization. Yellow fever situation report, 9 June2016. http://www.who.int/emergencies/yellow-fever/situation-reports/9-june-2016/en/

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

26. 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.
27. 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.
28. Sall AA, Faye O, Diallo M, Firth C, Kitchen A, Holmes EC. Yellow fever virus exhibits slower evolutionary dynamics than dengue virus. *J Virol*. 2010. 84(2):765-772.
29. Grobbelaar, A. A., Weyer, J., Moolla, N., Jansen van Vuren, P., Moises, F., & Paweska, J. T. Resurgence of Yellow Fever in Angola, 2015-2016. Emerging infectious diseases, 2016. *22*(10), 1854–1855.
30. Motayo B.O, Oluwasemowo O.O, Akinduti P.A. Evolutionary dynamics and geographic dispersal of beta coronaviruses in African bats. *PeerJ*. 2020;8:e10434. 26. doi:10.7717/peerj.10434.
31. Akinduti P.A, Ayodele O, Motayo BO, Obafemi YD, Isibor PO, Aboderin OW. Cluster analysis and geospatial mapping of antibiotic resistant *Escherichia coli* O157 in southwest Nigerian communities. *One Health*. 2022;15:100447. doi:10.1016/j.onehlt.2022.100447.
32. Akinduti P.A, Osiyemi JA, Banjo TT, et al. Clonal diversity and spatial dissemination of multi-antibiotics resistant Staphylococcus aureus pathotypes in Southwest Nigeria. *PLoS One*. 2021;16(2):e0247013. doi:10.1371/journal.pone.0247013
33. Li Y. Molecular epidemiology of yellow fever virus in Africa: A perspective of the phylogeographic split between East/Central African and West African lineages. Acta tropica, 2022. 225, 106199. https://doi.org/10.1016/j.actatropica.2021.106199.
34. Vasconcelos PF, Bryant JE, da Rosa TP, Tesh RB, Rodrigues SG, et al. Genetic divergence and dispersal of yellow fever virus, Brazil. Emerg Infect Dis. 2004. 10: 1578–1.
35. 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.
36. Hill, S. C., de Souza, R., Thézé, J., Claro, I., Aguiar, R. S., Abade, L., Santos, F. C. et al,. Genomic Surveillance of Yellow Fever Virus Epizootic in São Paulo, Brazil, 2016 - 2018. *PLoS pathogens*, 2020. *16*(8), e1008699. https://doi.org/10.1371/journal.ppat.1008699

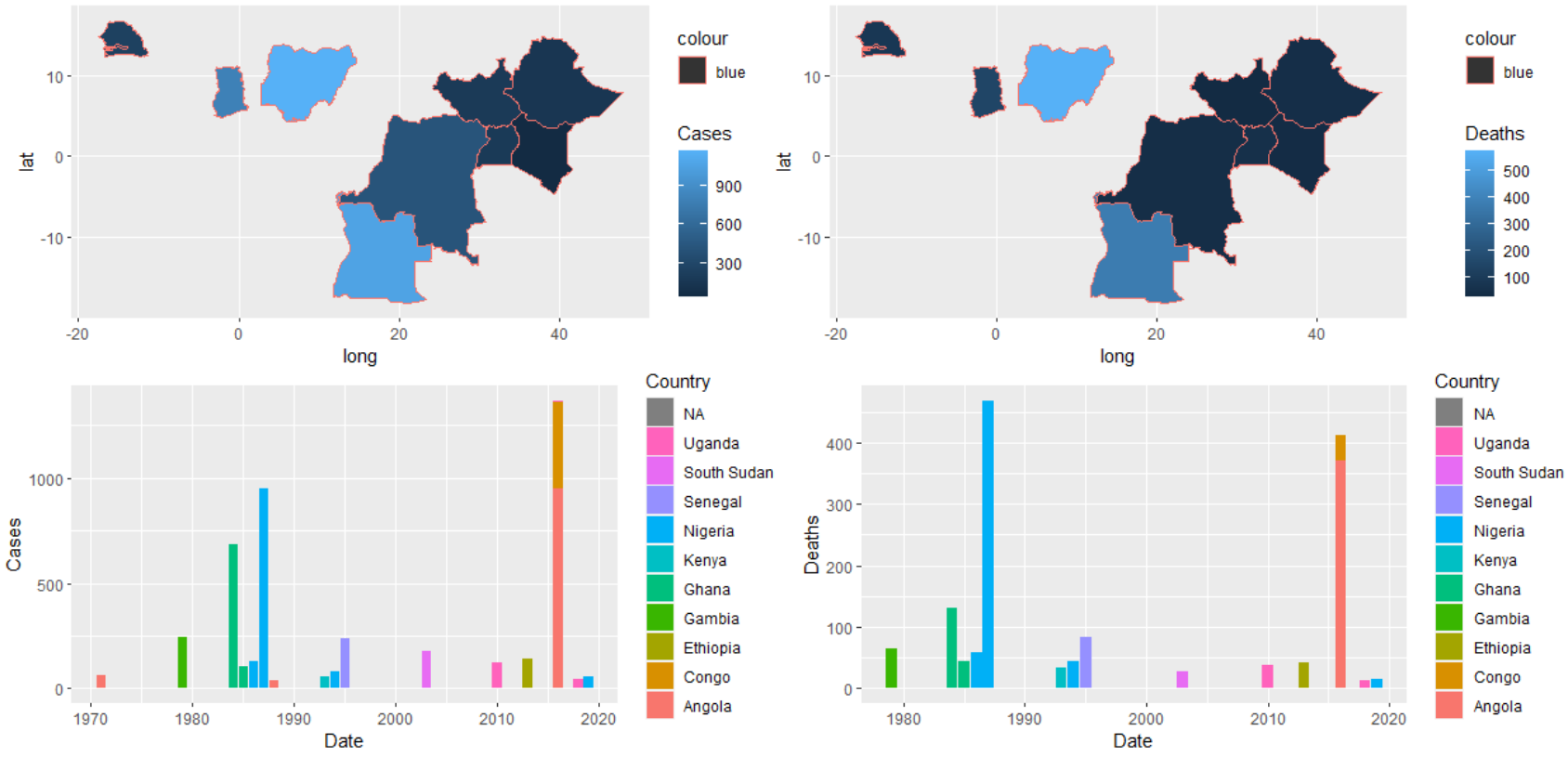


Figure 1 (top) Geographic distribution of YF virus cases, and deaths respectively across Africa, color intensity is representative of the number of confirmed YF cases in each country and is displayed as a color gradient by the side of each map. The NA represents not available. **1 (bottom)** Bar charts showing the number of confirmed Yellow Fever cases and deaths from 1970 to 2021 according to Country. The legend is color coded to represent the various countries where outbreaks occurred.

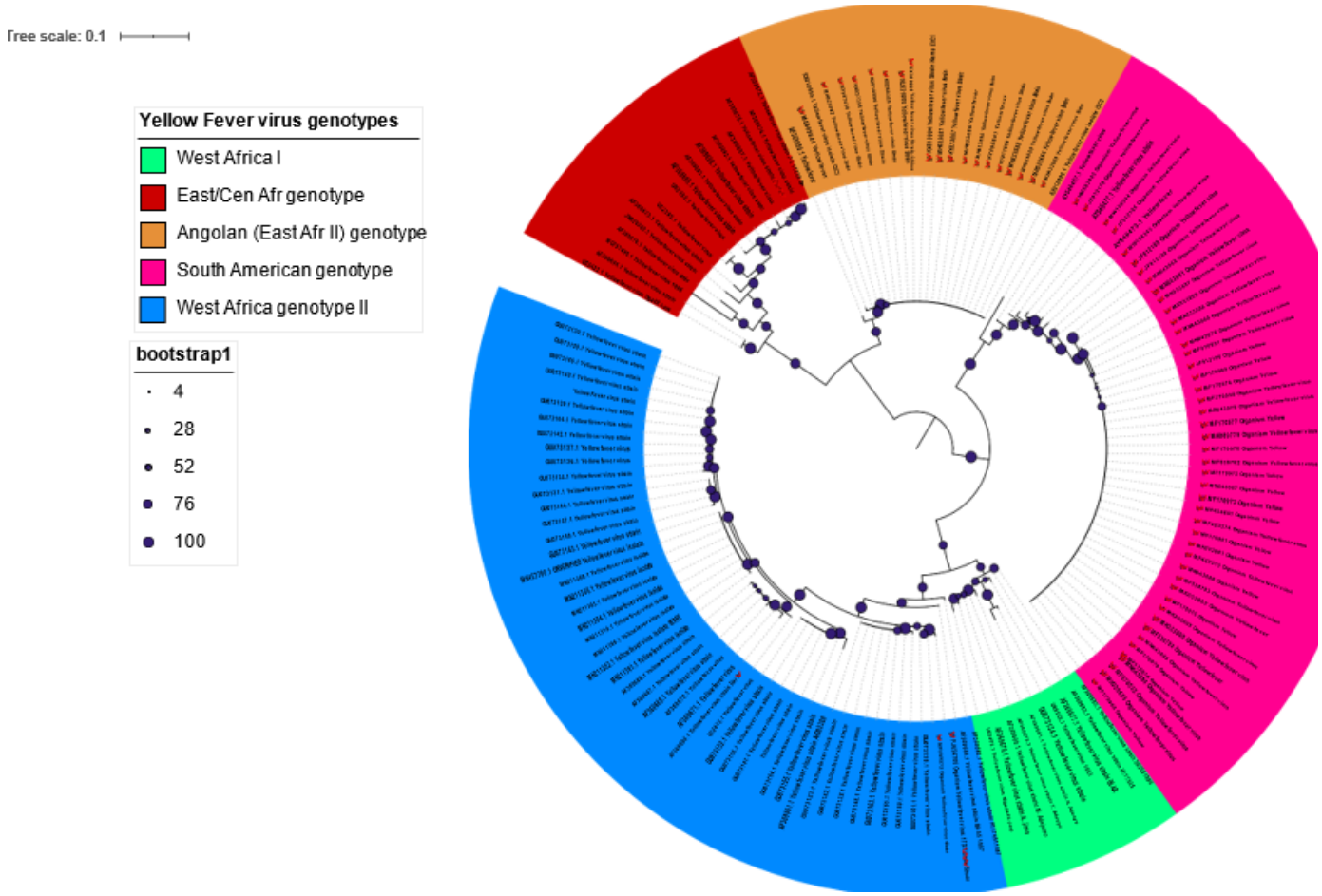


Figure 2. Maximum Likelyhood tree of African Yellow Fever isolates along with some global reference sequences.

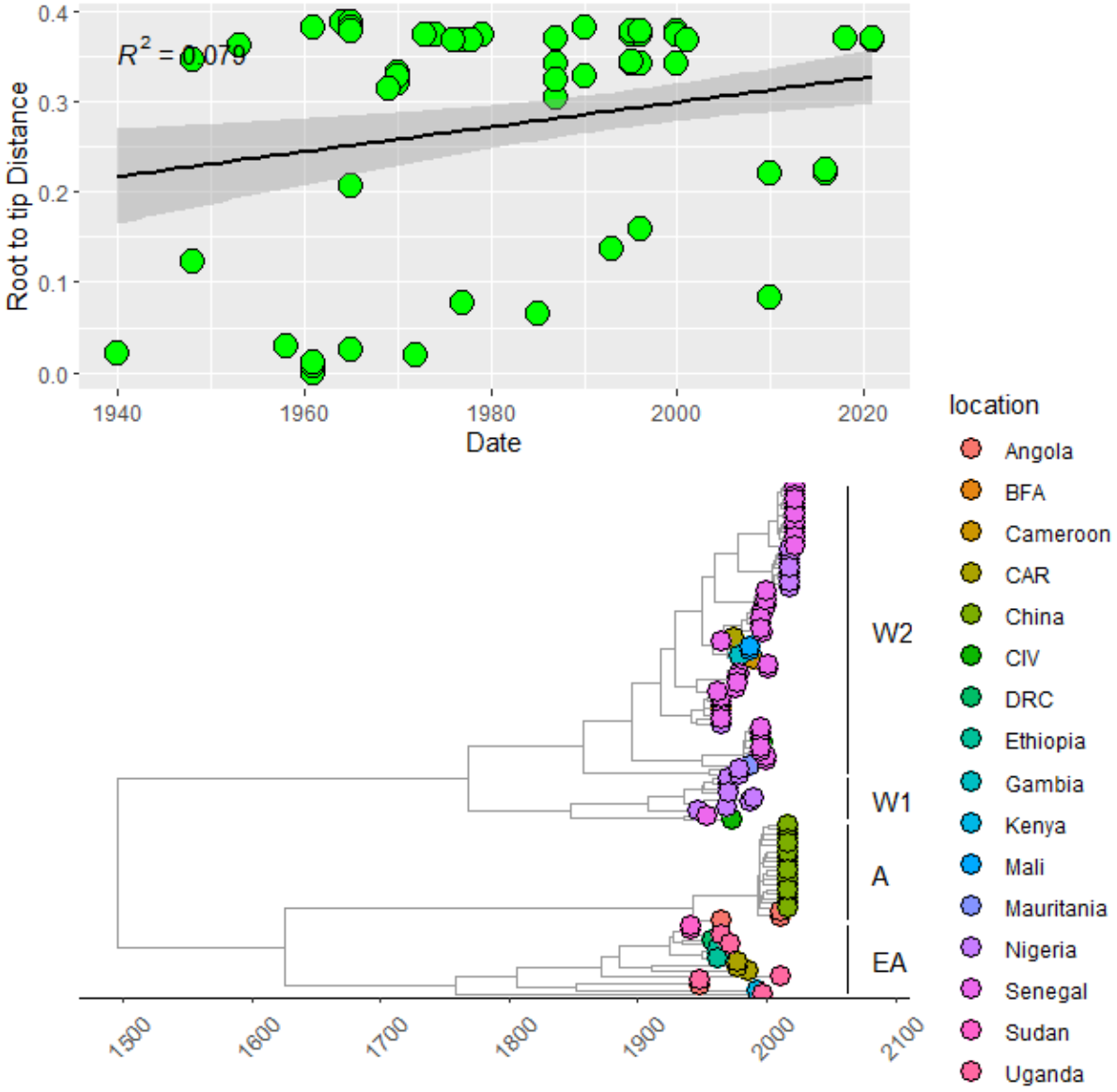
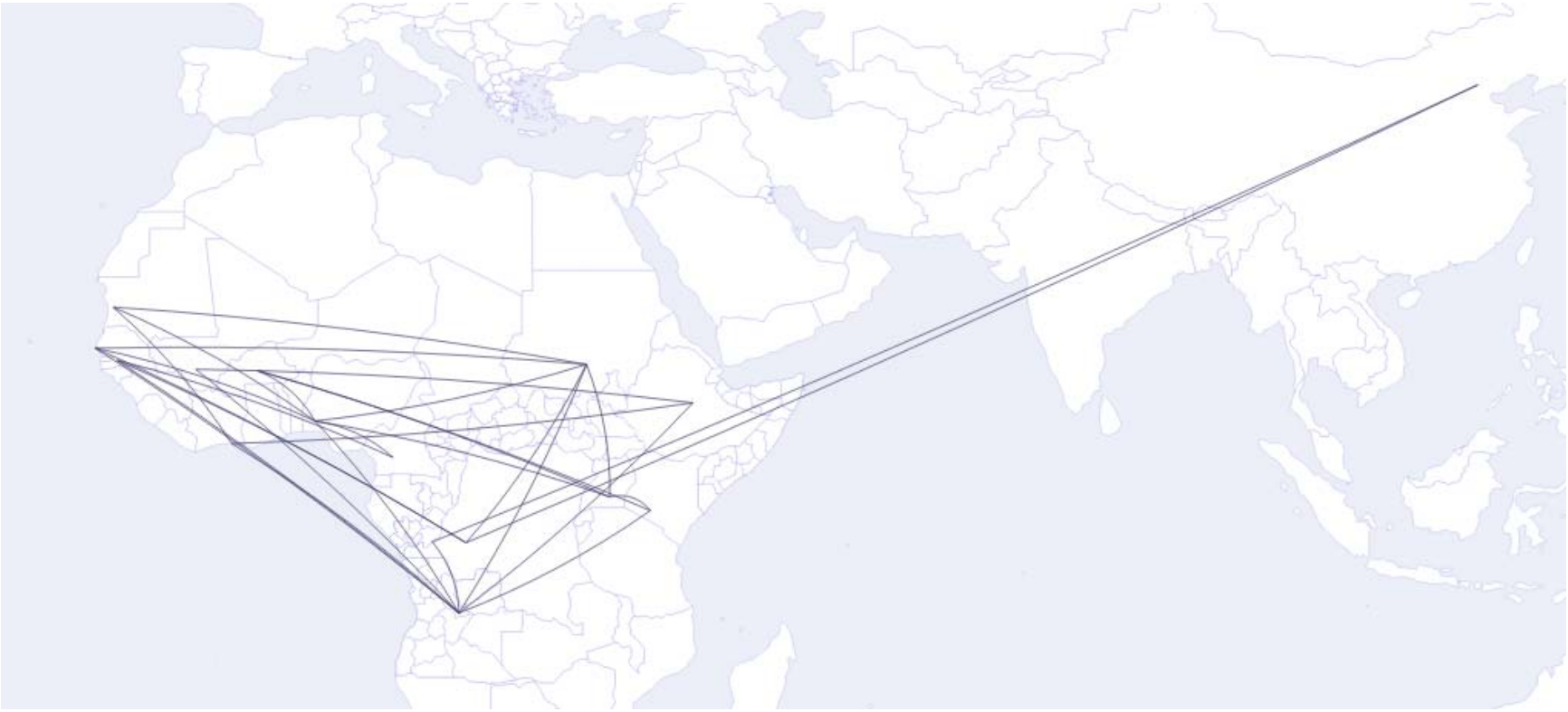


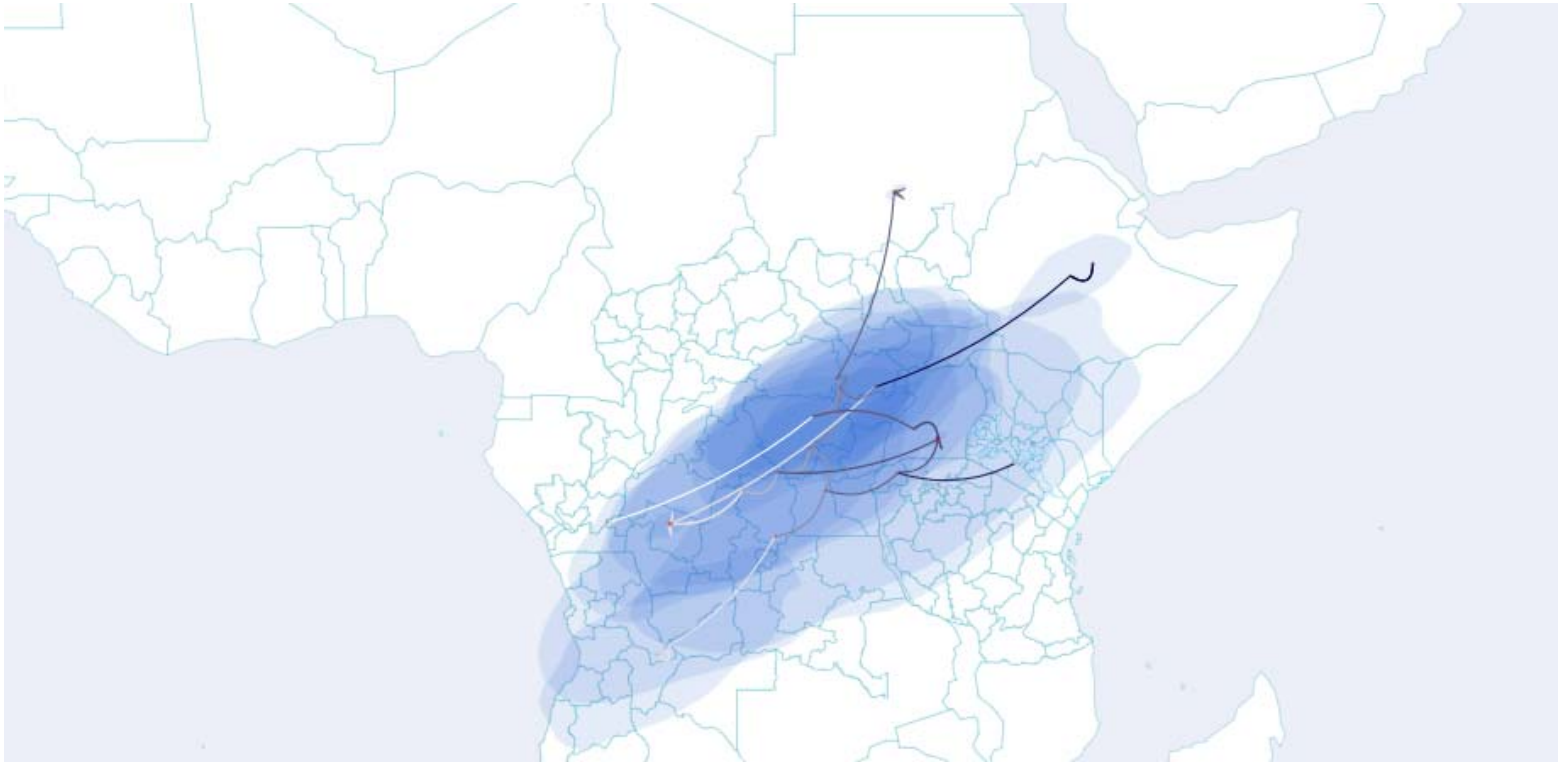
Figure 3 (top). Root to tip regression analysis of African Yellow Fever sequences. **3 (bottom)**. Time resolved Phylogeographic MCC tree of African Yellow virus E gene sequences, showing the various Countries of origin as tips. The genotypes are represented by horizontal bars by the side of the tree. WA represents West Africa, EA East Africa, and A Angola.



Bayes Factor = 0.3

Bayes Factor 0.5

Figure 4. Spatiotemporal reconstruction of the discrete phylogeography of Yellow Fever virus in Africa. Curved lines depict links between the nodes. The Bayes Factor is set at 0.3 (top) and 0.5 (bottom). Directionality of the movements is anticlockwise downward curve left to rightward movement; upward curve is right to leftward movement.



5a

5b

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, downward curve left to rightward movement; upward curve is right to leftward movement. **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, downward curve left to rightward movement; upward curve is right to leftward movement.