Virus Evolution, 2025, **11(1)**, veae116
DOI: https://doi.org/10.1093/ve/veae116
Advance Access Publication 30 December 2024
Research Article

Geographic origin and evolution of dengue virus serotypes 1 and 3 circulating in Africa

Sindiso Nyathi[1,*], Izabela M. Rezende[2], Katharine S. Walter[3], Panpim Thongsripong[4], Francis Mutuku[5], Bryson Ndenga[6], Joel O. Mbakaya[6], Gladys Agola[6], David M. Vu[7], Shannon Bennett[8], Erin A. Mordecai[9], Jason R. Andrews[2], A. Desiree LaBeaud[1,7]

[1]Department of Epidemiology and Population Health, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, United States
[2]Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, United States
[3]Division of Epidemiology, University of Utah, 295 Chipeta Way, Salt Lake City, UT 84108, United States
[4]Florida Medical Entomology Laboratory, Institute of Food and Agricultural Sciences, University of Florida, 200 9th St SE, Vero Beach, FL 32962, United States
[5]Department of Environmental and Health Sciences, Technical University of Mombasa, Mombasa, Kenya
[6]Center for Global Health Research, Kenya Medical Research Institute, P.O. Box: 1578 - 40100, Kisumu, Kenya
[7]Department of Pediatrics, Division of Infectious Diseases, Stanford University School of Medicine, Stanford University, 730 Welch Road, Stanford, CA 94305, United States
[8]Department of Microbiology, California Academy of Sciences, 55 Music Concourse Drive, San Francisco, CA 94118, United States
[9]Department of Biology, Stanford University, 371 Jane Stanford Way, Stanford, CA 94305, United States

*Corresponding author. Department of Epidemiology of Microbial Diseases, Yale School of Public Health, 60 College Street, New Haven, CT 06520, United States. E-mail: sindiso.nyathi@yale.edu

Abstract

Despite the increasing burden of dengue in Kenya and Africa, the introduction and expansion of the virus in the region remain poorly understood. The objective of this study is to examine the genetic diversity and evolutionary histories of dengue virus (DENV) serotypes 1 and 3 in Kenya and contextualize their circulation within circulation dynamics in the broader African region. Viral RNA was extracted from samples collected from a cohort of febrile patients recruited at clinical sites in Kenya from 2013 to 2022. Samples were tested by polymerase chain reaction (PCR) for DENV presence. Five DENV-positive samples were serotyped, and complete viral genomes for phylogenetic inference were obtained via sequencing on Illumina platforms. Sequences generated in our study were combined with global datasets of sequences, and Bayesian and maximum likelihood methods were used to infer phylogenetic trees and geographic patterns of spread with a focus on Kenya and Africa as a whole. Four new DENV-1 and one new DENV-3 genomes were successfully sequenced and combined with 328 DENV-1 and 395 DENV-3 genomes from elsewhere for phylogenetic analyses. The DENV-1 sequences from our study formed a monophyletic cluster with an inferred common ancestor in 2019 (most recent common ancestor 2019 and 95% high posterior density 2018–19), which was closely related to sequences from Tanzania. The single DENV-3 sequence clustered with sequences from Tanzania and Kenya, was collected between 2017 and 2019 and was related to recent outbreaks in the region. Phylogenetic trees resolved multiple clades of DENV-1 and DENV-3 concurrently circulating in Africa, introduced in the early-to mid-2000s. Three DENV-1 and four DENV-3 clades are highlighted, introduced between 2000 and 2015. Phylogeographic models suggest frequent, independent importations of DENV lineages into Kenya and Africa from East and South-East Asia via distinct geographic pathways. DENV-1 and DENV-3 evolutionary dynamics in Africa are characterized by the cocirculation of multiple recently introduced lineages. Circulating lineages are introduced via distinct geographic pathways that may be centered around regional nexus locations. Increased surveillance is required to identify key regional locations that drive spread, and dengue interventions should focus on interrupting spread at these locations.

Keywords: dengue virus; arboviruses; phylogeography; spread; Africa; Kenya

Introduction

Dengue virus (DENV) is an emerging arbovirus of great public health importance (Bhatt et al. 2013, Messina et al. 2014). The geographic expansion of the virus is linked to human and ecological factors such as climate change, large urban centers, and global travel, as well as evolutionary factors such as genetic diversity (Brady and Hay 2020). DENV outbreaks are being observed in Kenya and Africa more broadly with increasing frequency and are likely driven by both importations and local undetected spread (Vu et al. 2017, Gainor et al. 2022, Bosire et al. 2023). In order to respond to the threat of DENV in the African region, effective surveillance to understand the spread of the virus is required.

DENV exists as four genetically distinct serotypes (DENV-1, DENV-2, DENV-3, and DENV-4), which are further divided into genotypes (Rico-Hesse 1990). The genetic diversity of DENV results in complex epidemiological behaviors such as serotype/genotype replacement (Cummings et al. 2004, Zhang et al. 2005), wave-like dynamics, and geographic variations in detection and severity.

Several regions in the tropics are known to be endemic for at least one DENV serotype. Recently, multiple serotypes have been concurrently detected in some regions, suggesting hyperendemic circulation, which is characterized by the concurrent circulation of multiple serotypes within a location (Messina et al. 2014). Hyperendemicity contributes to a greater risk of severe disease due to the immunological phenomenon of antibody-dependent enhancement (Halstead 1988). As such, understanding the circulation dynamics of multiple dengue serotypes is a key element of DENV surveillance and outbreak response.

DENV has been circulating in Kenya and East Africa since the middle of the 20th century, and its spread is thought to have been accelerated by the increased travel associated with World War II (Messina et al. 2014, Brady and Hay 2020). Despite early detection in Kenya in the 1940s and an initial outbreak in 1982, little DENV sequence data exist between the time of initial detection and the early 2000s (Johnson et al. 1982a, 1982b). Serological studies in Kenya provide some evidence of past circulation of multiple serotypes in the region, but no molecular studies exist (Labeaud et al. 2007, Sutherland et al. 2011). The first confirmed DENV outbreak since 1982 was in Mombasa in 2011, and since then several outbreaks have been detected in coastal towns and inland cities such as Kisumu (Obonyo et al. 2018). Due to this relative lack of data from early outbreaks, early dynamics of DENV circulation, including any serotype dominance or replacement events that occurred, remain unknown. Recent work indicates that although all four serotypes occur during epidemic and interepidemic periods, DENV-2 is the most prevalent (Shah et al. 2020, Pollett et al. 2021). While previous work has shown that DENV-2 circulation in Kenya is characterized by multiple cocirculating lineages, likely independently introduced (Nyathi et al. 2024), the introduction and expansion of the additional circulating DENV-1 and DENV-3 serotypes have not been contextualized within the circulation dynamics in the broader African region.

In this study, we investigate the genetic diversity of DENV-1 and DENV-3 serotypes in Kenya relative to the broader African region. Specifically, we ask if DENV-1 and DENV-3 circulation and outbreaks in the region are driven by low-level undetected circulation or importations. We examine the local and regional spread of DENV-1 and DENV-3 serotypes and determine the geographic histories of detected lineages. Finally, we test whether the geographic routes of DENV-1 and DENV-3 circulation in Africa are two realizations of the same underlying stochastic transmission process or, alternatively, governed by distinct underlying transmission processes. Understanding the circulation and expansion of DENV in Africa will enable more effective disease control, outbreak response, and preparedness.

Methods

Clinical surveillance

Data were collected as part of two arbovirus surveillance cohorts in Kenya, Cohort A and Cohort B. The cohorts have been characterized in detail in previous work (Shah et al. 2020, Vu et al. 2023, Kiener et al. 2024, Nyathi et al. 2024, Tariq et al. 2024). Briefly, for both cohorts, participants were recruited into the study from outpatient clinics in one of four dengue endemic sites in Kenya: Chulaimbo, Kisumu, Msambweni, and Ukunda as shown in Fig. 1. For Cohort A, children (1–17 years) who presented with an acute febrile illness (reported during the previous 2 weeks and including temperature ≥38°C at presentation) were recruited into the study from all four study sites between January 2014 and June 2018. For

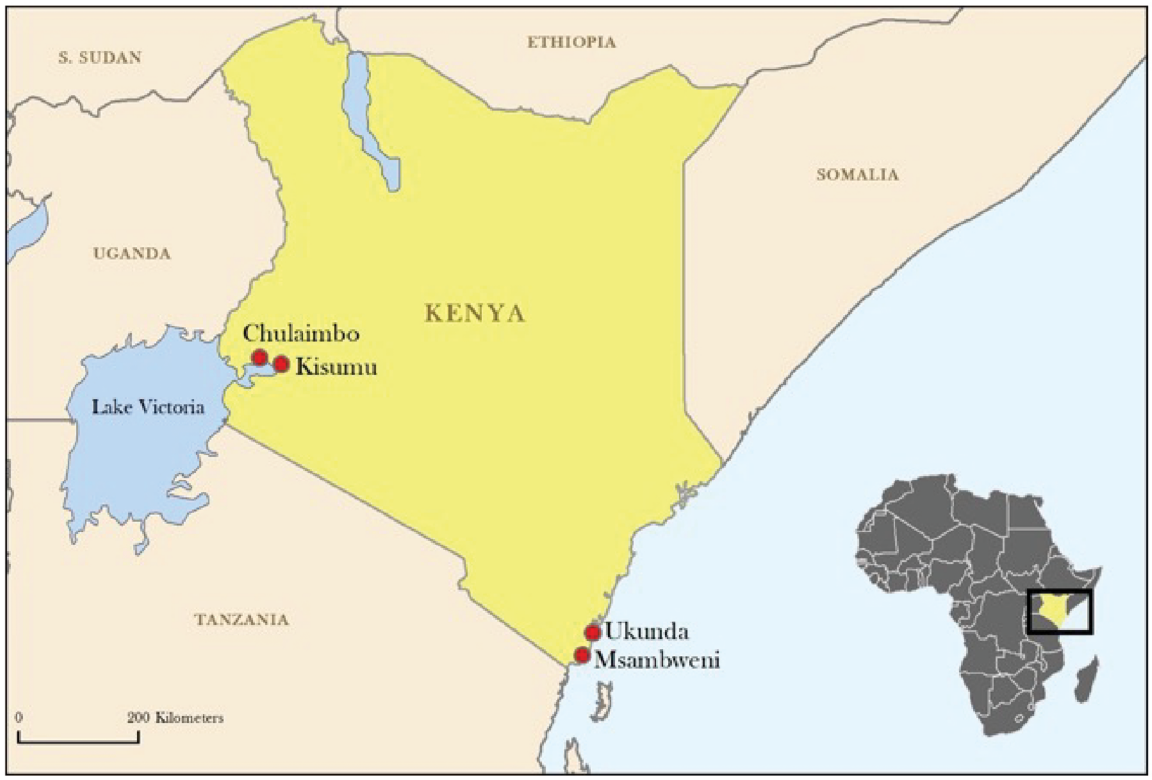

Figure 1. Locations of four data collection sites in Kenya. Participants were recruited from outpatient clinics in Chulaimbo, a rural inland site; Kisumu an urban inland site; Msambweni, a rural coastal site; and Ukunda, an urban coastal site.

Cohort B, children and adults from the Ukunda and Kisumu sites who were part of an ongoing longitudinal cohort were followed through clinical visits during acute febrile illness events between July 2018 and June 2022. A comprehensive clinical examination was conducted including collection of demographic and clinical history data and blood draw for serological and molecular testing. Demographic data were stored in RedCap (Version 12.2.11) hosted at Stanford University, while human samples were frozen, shipped to Stanford University, and stored at −70°C until use (Harris et al. 2009). Ethical approval and oversight for data collection for this study were obtained from the Institutional Review Board of Stanford University (IRB 31488), as well as the Kenya Medical Research Institutes (KEMRI SSC 2611) and Technical University of Mombasa Ethical Review Committee (TUM/ERC EXT/004/2019).

Molecular testing and DENV-1 and DENV-3 viral amplification

Nucleic acid material was extracted from total blood using RNA extraction kits (Omega) following the manufacturer's instructions. Total RNA was converted into complementary DNA (cDNA) and tested for viral presence using a SuperScript III One-Step qPCR assay with a set of published primers and probes capable of detecting all four DENV serotypes (Waggoner et al. 2016). DENV-positive samples were serotyped via an additional quantitative reverse transcription PCR (RT-qPCR), with a set of published serotype-specific primers and probes (Waggoner et al. 2013). DENV-positive samples identified in the serotype-specific RT-qPCR reaction were run through a viral amplicon enrichment protocol using custom, serotype-specific primers. Custom DENV enrichment primers were designed to capture the circulating diversity of each DENV serotype circulating in Africa using PrimalScheme (Supplementary Material Appendix A) (Grubaugh et al. 2019). Primers were designed to enrich the genome in 400-bp amplicons with an average overlap of 75 bp between amplicons.

Next-generation sequencing and bioinformatics pipeline

Purified, enriched cDNA was cleaned using AMPure XP beads (Beckman Coulter) and quantified using a Qubit 4 assay (Thermo Fisher Scientific), and library preparation was conducted using the SparQ library preparation kit (Quantabio) with Illumina TruSeq adapters. Additional quality control steps were conducted using the Agilent Bioanalyzer assay to determine average fragment size and the KAPA HyperPrep kit (Roche Sequencing) to determine final concentrations. The prepared samples were pooled in equimolar amounts and sequenced on Illumina's MiSeq platform using a 2 × 250-bp V2 kit and paired-end dual-indexed sequencing. Blanks and no template controls were included in all reactions from RNA extraction to sequencing.

Sequencing data were processed using custom scripts based on the iVar bioinformatics pipeline (Grubaugh et al. 2019). Raw reads were downloaded from Illumina's BaseSpace platform and aligned to a serotype-specific reference downloaded from GenBank (DENV-1—MT076932 and DENV-3—MT076948) using *BWA-MEM* (Li and Durbin 2010). *Fastqc* was used to visualize read quality data. Adapters and enrichment primer sequences were trimmed from aligned reads using the *iVar trim* command, and low-quality reads were trimmed using a PHRED score cutoff of 20. Reads whose length was <10 bases after trimming and quality filtering were discarded. Trimmed reads were sorted using *samtools*, and consensus genomes were generated using *samtools mpileup* and *ivar consensus* commands (Li et al. 2009). The consensus base at a position was called for a minimum depth of 10 reads and at a frequency of at least 80%. *samtools* was used to retrieve coverage and depth statistics. Consensus genomes for the sequences generated as part of this study were exported for phylogenetic analysis and also uploaded to GenBank with accession numbers as shown in Supplementary Table S1.

Phylogenetic analyses

Consensus genomes generated from this study were combined with publicly available complete or partially complete (envelope gene >1400 bp) DENV-1 and DENV-3 genomes downloaded from GenBank. A 1400-bp cutoff was used for inclusion in order to include several DENV genomes sequenced at the envelope gene level from locations in Africa. Sequence collection date and geographic location data were also downloaded from GenBank. For countries with many genomes available, a hierarchical clustering approach was used on preliminary country-specific datasets to define clusters of high similarity from the same location based on distance-based similarity measures and randomly sample a fraction of the sequences from within clusters to trim the sequence dataset. To do this, geographic location information from countries with sequence and location data available was used to create location-specific sequence datasets. Location-specific sequence datasets were separately aligned and hierarchically clustered based on similarity evaluated with pairwise distance-based measures using the Bioconductor package, MSA2dist (Ullrich 2024). Similar clusters in each location were randomly down-sampled by a factor r, depending on the original location-specific dataset size, N (i.e. $N < 30$, $r = 0.25$; $30 < N < 100$, $r = 0.125$; $100 < N < 500$, $r = 0.067$; and $N > 500$, $r = 0.05$). This served to reduce the size of the dataset while maintaining information about the geographic spread contained in sequences. Sequences in the final dataset were genotyped using Genome Detective's Dengue Typing tool (https://www.genomedetective.com/app/typingtool/dengue/). TempEst (version 1.5.3) was used to evaluate the suitability of our data for timed-phylogeny analyses with root-to-tip regressions (Supplementary Fig. S1a and b) (Rambaut et al. 2016, Fonseca et al. 2019). Outliers as indicated by the TempEst residuals were excluded from the analysis (Chen et al. 2021).

The sequence datasets for each serotype were aligned using MAFFT (version 7.310), and the alignments were examined and manually curated in AliView, including pruning of redundant sequences as well as the exclusion of 3′ and 5′ untranslated regions (Larsson 2014). RDP4 was used for recombination detection, and additional sequences with evidence of recombination were excluded from the analysis (Supplementary Materials Appendix B).

Bayesian and maximum likelihood (ML) approaches were used to infer phylogenetic trees. Initial phylogenetic trees were estimated using an ML approach with RAxML-NG (Kozlov et al. 2019). ModelTest-NG (version 0.1.7) was used to determine the best nucleotide substitution model on the ML tree (Supplementary Tables S2 and S3) (Posada and Crandall 1998, Darriba et al. 2020). Bayesian tree estimation was conducted on BEAST 1.10.4 using a relaxed molecular clock and the general time reversible (GTR) nucleotide substitution model with a Gamma distributed rate and invariant sites (GTR + G4 + I), as determined by ModelTest-NG. Three independent runs of a 150 million step Markov Chain Monte Carlo (MCMC) procedure were conducted, with sampling every 15 000 generations and 10% burn-in. Output convergence and effective sample sizes were evaluated using Tracer (version 1.7), with independent runs down-sampled and combined in LogCombiner and final tree data generated in TreeAnnotator (version

1.10). Maximum clade credibility (MCC) phylogenetic trees were visualized and modified using FigTree (version 1.4) and the R *ggtree* package (version 3.8.2). Key clades corresponding to DENV-1 and DENV-3 circulation in Kenya and Africa are highlighted. These "clades" of interest were defined as groups of sequences within the genotypes descended from a single ancestral introduction into a defined study region in Africa (i.e. Northern Africa, Southern Africa, Eastern Africa, Central Africa, or Western Africa) and exhibiting continued general circulation in Africa. Evolutionary characteristics including evolution rates and time to most recent common ancestors (tMRCAs) were calculated based on the nodes of the MCC tree, and uncertainty was expressed using 95% high posterior density (HPD) intervals.

Phylogeographic analyses

Phylogeography of DENV-1 and DENV-3 in Africa was jointly estimated along with the timed tree, using a Bayesian Stochastic Search Variable (BSSV) selection procedure to evaluate support, with location included as a discrete trait and an asymmetric substitution model to account for different rates of inward versus outward gene flow and ancestral state reconstruction of the location trait. Sequence country data were recategorized into 16 regions to create a more computationally feasible dataset (Supplementary Table S6), with samples collected from geographic locations in Africa grouped into Southern, Western, Central, Eastern, or Northern Africa. To quantify the contribution of key regions to the genetic diversity of DENV in Kenya and Africa, the number of Markov jumps from one location to another and the reward time spent in each location, across the history of the inferred phylogeny, were stored and visualized, along with geographic transitions across regions of interest. Finally, to investigate the cocirculation of DENV-1 and DENV-3 clades and evaluate the potential for hyperendemicity in study regions within the African continent, the Markov jump histories for each clade were summarized and combined. To summarize the "reward times" or periods of time, a lineage remains in a given location, for each location, a clade was inferred to have circulated in, and the median start and end times from the Markov jump history were calculated for each location for that clade and visually inspected.

Comparing DENV-1 and DENV-3 geographic transmission routes

To evaluate whether the reported DENV-1 and DENV-3 transmission routes result from the same underlying stochastic process, or from distinct underlying processes, we used path-sampling/stepping-stone sampling approaches to estimate the marginal likelihoods of the two models. To do this, the DENV-1 and DENV-3 sequence alignments were included as separate partitions in the same .xml file. Bayesian tree estimation and phylogeography were jointly conducted on each partition, with a relaxed molecular clock and the GTR nucleotide substitution model with a Gamma distributed rate and invariant sites (GTR + G4 + I), which were unlinked for the two partitions to account for underlying differences in nucleotide substitution or molecular clock processes. The same 16 regions as in the primary analysis were used to define serotype-specific location traits (i.e. DENV-1 location trait and DENV-3 location trait), a Bayesian stochastic search variable procedure was used to evaluate support for transition rates, and ancestral states were reconstructed. While the clock models were linked for both serotype-specific location partitions, in order to test whether DENV-1 and DENV-3 transmission could be described by a single rate matrix, we ran a model in which the substitution models for the location partitions were linked to represent a single underlying process (M_1), and another in which they were unlinked to represent two separate underlying processes (M_2). Both models were run for 100 million steps of the MCMC procedure, with sampling every 1000 steps and 10% burn-in. The path-sampling/stepping-stone sampling procedures to estimate the marginal likelihoods of both models were run for 100 steps, each with a chain length of 1 million and sampling every 1000 MCMC steps within each chain. Both marginal likelihood values and their difference, representing the log Bayes Factor, are reported to compare models.

Posterior node support values are reported to indicate support for each clade or branch in the MCC trees. Location probability values are also reported for each branch to indicate the probability of the ancestral location inferred from the phylogeographic analysis. Auxiliary data analysis and final data visualization were conducted in R (version 4.1.2). The sequences generated as part of this study have been deposited to GenBank with accession numbers: OR765987–OR765990 and PQ412806 (https://www.ncbi.nlm.nih.gov/genbank/). All other sequences used in this analysis are available in the Supplementary data (Supplementary Tables S4 and S5). All .xml files generated as part of these analyses are available online (https://github.com/SindisoNyathi/doctoralrepo_dengue_phylo/tree/main/beast_xmls).

Results

A total of four DENV-1 (Fig 2) and one DENV-3-positive (Fig 3 samples in this study had adequate sequencing coverage and read quality (PHRED score > 20 and depth > 10 reads/position) for inclusion in downstream workflows, from a total of 31 DENV cases detected between 2019 and 2022 by PCR. Sequences from this study were combined with 328 DENV-1 and 395 DENV-3 sequences downloaded from GenBank and collected from elsewhere to create serotype-specific datasets. The predominant genotypes in the study datasets were DENV-1 Genotype III (93 of 332 sequences) and DENV-3 Genotype III (249 of 396 sequences) (Supplementary Tables S4 and S5). The four DENV-1 samples collected in this study fell into DENV-1 Genotype III, while the single DENV-3 sample clustered within DENV-3 Genotype III. The accession numbers of all sequences analyzed in this study, including collection date, location, and genotyping information, are included in the Supplementary data (Supplementary Table S1). Final datasets exhibited a strong temporal signal with the R^2 values of 0.70 for the DENV-1 dataset and 0.83 for the DENV-3 dataset (Supplementary Fig. S1). Nucleotide substitution rates resolved from Bayesian timed trees were relatively similar for both serotypes (DENV-1 7.17×10^{-4}, 95% HPD $6.57–7.77 \times 10^{-4}$ and DENV-3 9.49×10^{-4}, 95% HPD $8.74–10.17 \times 10^{-4}$).

ML and Bayesian trees resolved three recent clades related to DENV-1 circulating in Africa (Clades a, b, and c in Figs 2 and 4). The most recent was a DENV-1 Genotype III clade containing sequences collected from East Africa, in Tanzania between 2020–19 and in Kenya in 2020 (Figs 2 and 4, Clade a). Phylogeographic models suggest that this clade may have been introduced into East Africa from Upper Southeast Asia in 2009 (node date = 2009, 95% HPD 2007–11, and location probability = 0.86) (Fig. 4, Clade a). The DENV-1 sequences collected from our study sites formed a monophyletic clade within this larger clade (MRCA = 2019, 95% HPD 2018–19) (Fig. 4). The second clade was also in DENV-1 Genotype III and included sequences collected from Western Africa, including Burkina Faso, Cote d'Ivoire, and Senegal, from 2017 to 2019 and sequences collected from East Africa (Seychelles, Kenya, and Eritrea) in 2016 (Figs 2 and 4,

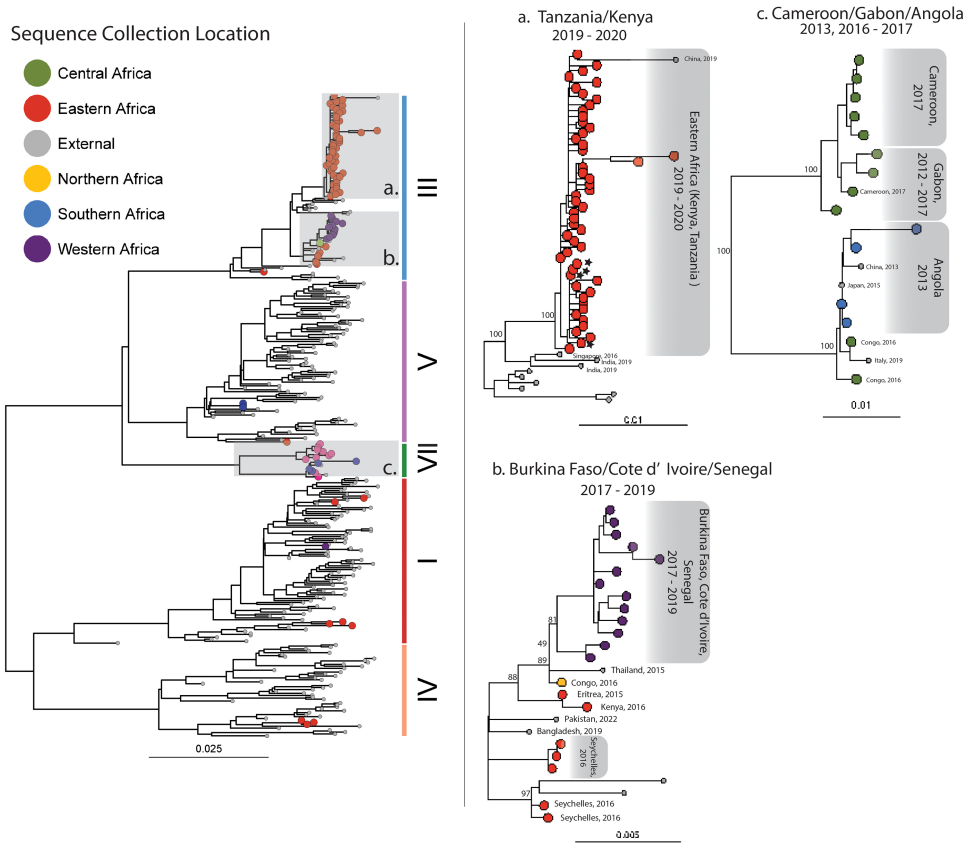

Figure 2. ML tree for DENV-1 showing three significant DENV-1 clades detected in Kenya and Africa inferred using RAxML-NG. Four DENV-1 sequences collected in this study were combined with 328 sequences from elsewhere, and RAxML-NG was used to infer ML phylogenetic trees. Tip colors indicate the region of sequence collection with regions in Africa shown in color as indicated in the legend and regions outside Africa shown in gray. Black stars indicate isolates from this study. Colored vertical bars on the main tree (left tree) indicate the genotype of DENV sequences in that clade. Significant clades of sequences collected in Africa are highlighted and shown as insets (Boxes a, b, and c) with country locations and collection dates of sequences in these clades labeled.

Clade b). This clade may have been introduced into Eastern Africa from Upper Southeast Asia in 2010 (node date = 2010, 95% HPD 2009–12, and location probability = 0.81) and eventually made its way into Western Africa via Central Southeast Asia in 2013 (node date = 2013, 95% HPD 2012–14, and location probability = 0.42) (Fig. 4). The oldest DENV-1 clade contained sequences from Central Africa (Cameroon and Gabon) collected from 2017 to 2021 and Southern Africa (Angola) collected in 2013 (Figs 2 and 4, Clade c). This clade was introduced from Central America/Caribbean in 1950 (node date = 1950, 95% HPD 1935–61, and location probability = 0.91). Additional smaller clusters of sequences collected from regions in Africa fell into DENV-1 Genotypes I, IV, and V and were collected from East Africa in 2004–06 and 2010–13 and Southern Africa in 1988 with geographic origins in Lower Southeastern Asia, Western Asia, and Central America/Caribbean, respectively (Supplementary Fig. S2).

Phylogenetic trees resolved four clades of DENV-3 circulating in Africa, with shared evolutionary history (Clades a, b, c, and d in Figs 3 and 5). The newest clade contains sequences from Western Africa (Senegal and Burkina Faso) collected from 2017 to 2019 and Central Africa (Cameroon) collected in 2017 (Figs 3 and 5, Clade a). This lineage was introduced from Western Asia (node date = 2003, 95% HPD 2002–04, and location probability = 0.92). The second clade contained sequences from Central Africa (Gabon) collected from 2016 to 2017 (Figs 3 and 5, Clade b). This clade was introduced from Western Africa (node date = 2006, 95% HPD 2005–07, and location probability = 1) and had recent shared ancestry with DENV-3 Clade a above. The third clade contained sequences from Eastern Africa (Djibouti, Comoros, and Madagascar) collected from 2010 to 2012 (Figs 3 and 5, Clade c) and introduced from Western Africa (node date = 2004, 95% HPD 2003–05, and location probability = 0.59). The final clade also contained sequences from East Africa (Kenya and Tanzania) collected from 2017 to 2019 (Figs 3 and 5, Lineage d), and may have been introduced from Upper Southeast Asia (node date = 2007, 95% HPD 2004–09, and location probability = 1). All four clades of DENV-3 are closely related and inferred to be descended from sequences from Upper Southeast Asia, imported into Africa independently in the early to mid-2000s.

Geographic transition rates among regions and support for each rate were inferred via a Bayesian stochastic search variable procedure. Transition rate matrixes corresponding to DENV-1 and DENV-3 circulation show high rates of gene flow out of Upper, Central, and Lower Southeast Asia, with a relatively high rate of gene flow for DENV-1 from Upper Southeast Asia into Eastern Africa and high rates of gene flow into East Asia (Supplementary Figs S5 and S6 Transition matrixes also show high rates of bidirectional gene flow between Central America and the Caribbean, and South America. Markov jump history persistence traces show that all the clades highlighted in this study, with the exception of DENV-1 Clade c, were introduced relatively recently and cocirculated for short periods of time before detection (Figs 6 and 7). DENV-1

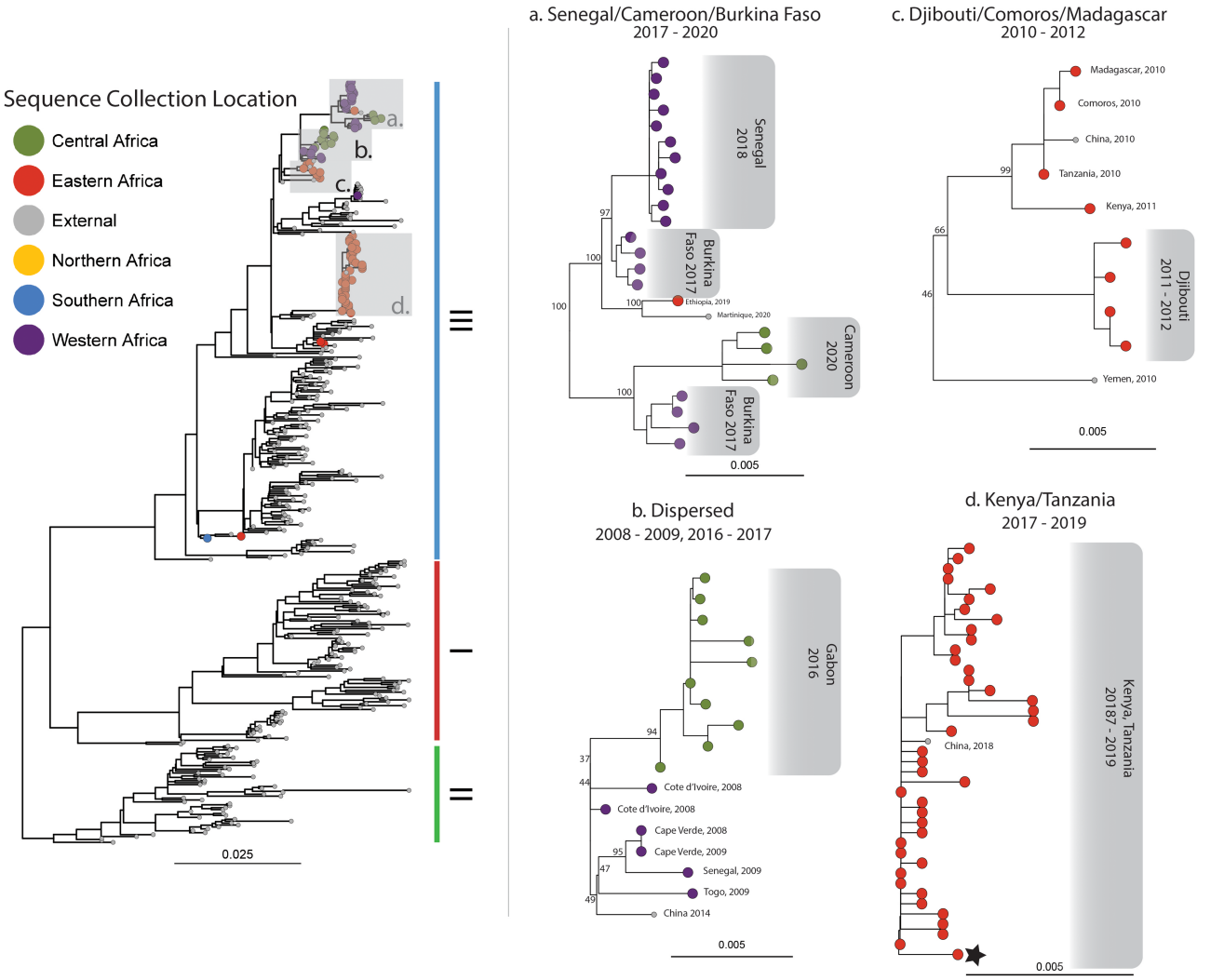

Figure 3. ML tree for DENV-3 showing four significant DENV-3 clades detected in Kenya and Africa inferred using RAxML-NG. One DENV-3 sequence collected in this study was combined with 395 sequences from elsewhere and RAxML-NG was used to infer ML phylogenetic trees. Tip colors indicate the region of sequence collection with regions in Africa shown in color as indicated in the legend and regions outside Africa shown in gray. The black star indicates the isolate from this study. Colored vertical bars on the main tree indicate the genotype of sequences in that clade. Four significant clades of sequences collected in Africa are highlighted and plotted as insets (Boxes a, b, c, and d) with country locations and collection dates of sequences in these clades labeled.

Clade c, unlike the other highlighted clades, exhibited prolonged circulation in Southern Africa before eventual detection in Central Africa (Fig. 6c). Western Africa and Eastern Africa contribute the most to introduction of clades into Africa and their subsequent spread to other locations and also exhibit cocirculation of DENV-1 and DENV-3 (Fig. 8).

Finally, in the comparison of two models of DENV-1 and DENV-3 circulation, M_1, the model in which DENV-1 and DENV-3 circulation routes are two realizations of the same underlying process had a log marginal likelihood of −196 454.63 estimated by path sampling and −196 523.75 estimated by stepping-stone sampling. M_2, the model in which DENV-1 and DENV-3 circulation routes are governed by two separate processes, had a log marginal likelihood of −196 470.50 estimated by path sampling and −196 540.13 estimated by stepping-stone sampling. These log marginal likelihood estimates give differences, denoting the log of the Bayes Factor, $\log(BF_{10}) = [\log \text{marginal likelihood } (M_1) - \log \text{marginal likelihood } (M_2)]$ of 15.88 from path sampling comparison and 16.38 from the stepping-stone sampling comparison in favor of M_1, the model in which DENV-1 and DENV-3 circulation routes are two realizations of the same underlying process and can thus be modeled using a single rate transition matrix.

Discussion

DENV has a unique ecology and epidemiology due to the emergence and subsequent spread of four genetically and antigenically distinct serotypes. In Kenya, serological and molecular evidence suggests hyperendemic circulation of DENV serotypes, causing frequent outbreaks in cities such as Malindi, Mombasa, and Kisumu (Ellis et al. 2015, Konongoi et al. 2016, Vu et al. 2017, Langat et al. 2020, Shah et al. 2020, Muthanje et al. 2022). Despite the initial detection of DENV in Kenya in 1982, limited molecular data exist examining the spread and circulation of DENV, particularly serotypes 1 and 3. DENV-1 and DENV-3 have also been detected in other countries in East Africa and Africa more broadly, causing recent outbreaks in Tanzania (Kelly et al. 2023), Burkina Faso (Letizia et al. 2022), Senegal (Faye et al. 2014, Gaye et al. 2021),

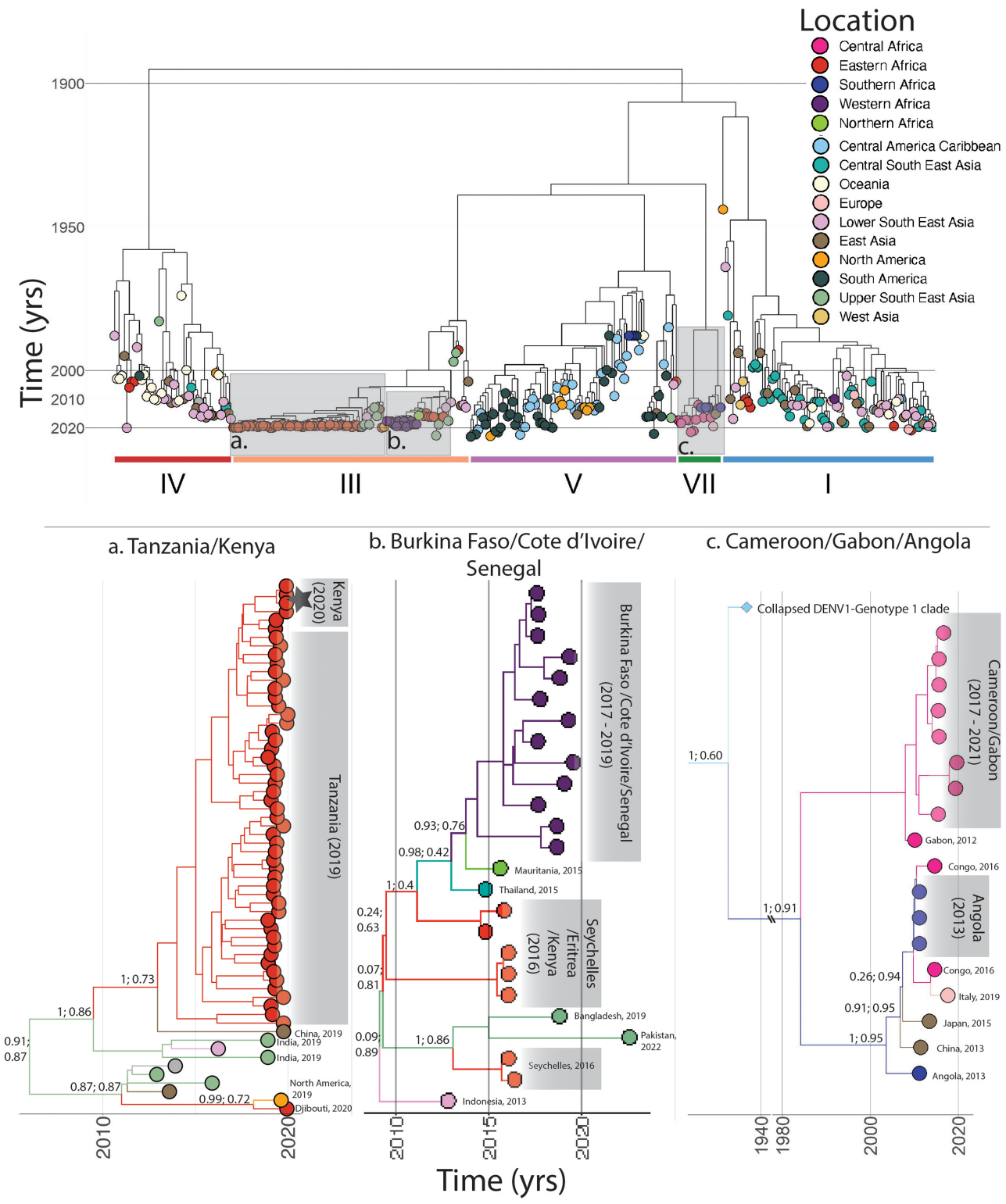

Figure 4. Timed Bayesian phylogenetic trees of DENV-1. Timed Bayesian phylogenetic trees and phylogeography were jointly inferred using the four DENV-1 isolates from this study and 328 publicly available sequences downloaded from Genbank. Branch colors on the inset trees correspond to inferred ancestral locations and tip colors indicate the isolate collection location, while colored bars indicate genotypes from the dengue typing tool. Three significant clades of sequences collected from Africa are highlighted (Boxes a, b, and c) and shown as insets. Collection location and dates of sequences in these clades are labeled within insets to highlight the geographic distributions of these clades. The black star indicates the clade of isolates from this study. Clades of interest are annotated with posterior node support values and location probability values (i.e. node support and location probability).

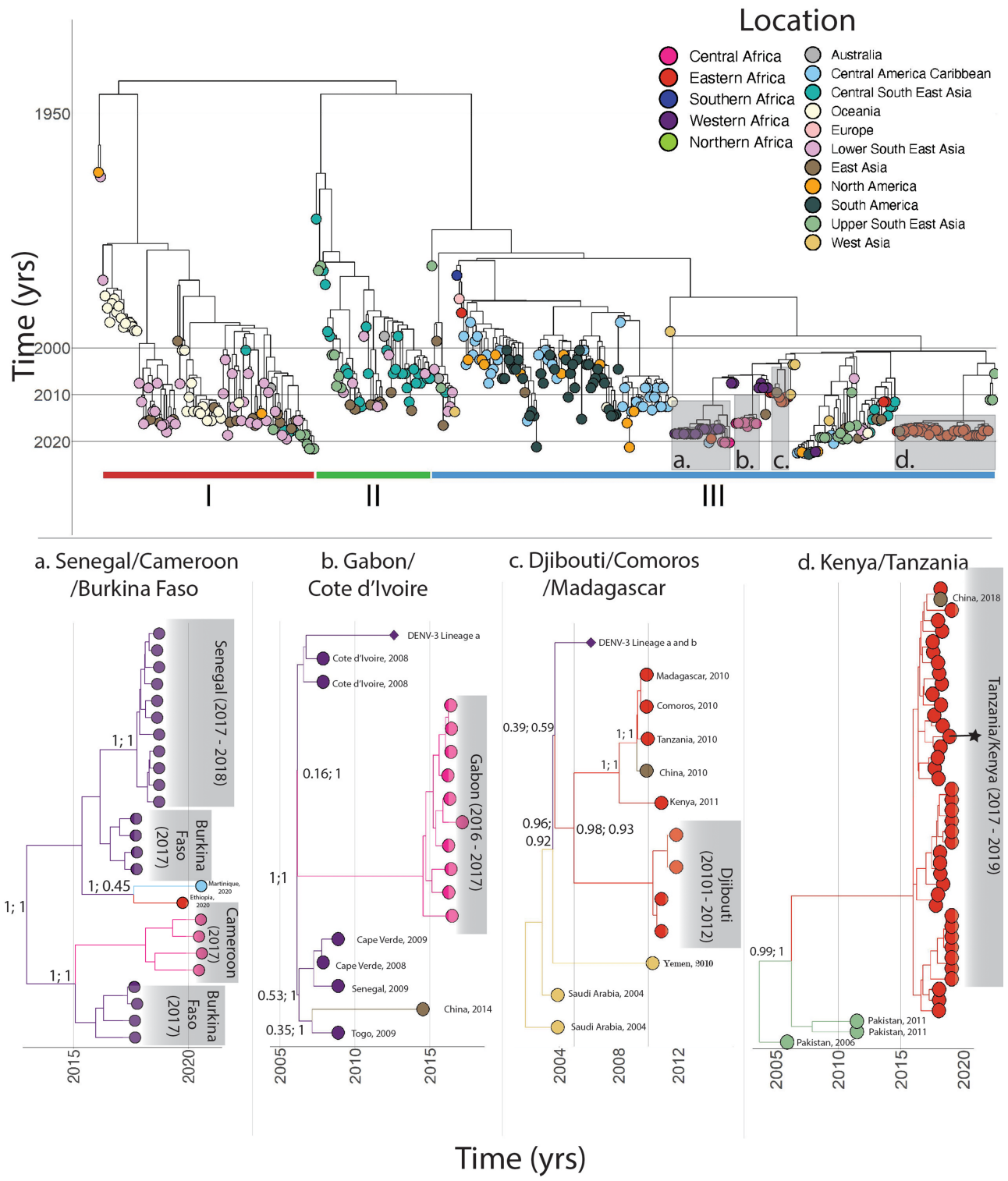

Figure 5. Timed Bayesian phylogenetic trees of DENV-3. Timed Bayesian phylogenetic trees and phylogeography were jointly inferred using one DENV-3 isolate from this study and 395 publicly available sequences downloaded from Genbank. Branch colors on the inset trees correspond to the inferred ancestral location and tip colors indicate the isolates collection location, while colored bars indicate genotypes from the dengue typing tool. Four significant clades of sequences collected from Africa are highlighted (Boxes a, b, c, and d) and shown as insets. Collection location and dates of sequences in these clades are labeled within insets to highlight geographic distributions of these clades. The black star indicates the sequence collected from this study. Clades of interest are annotated with posterior node support values and location probability values (i.e. node support and location probability).

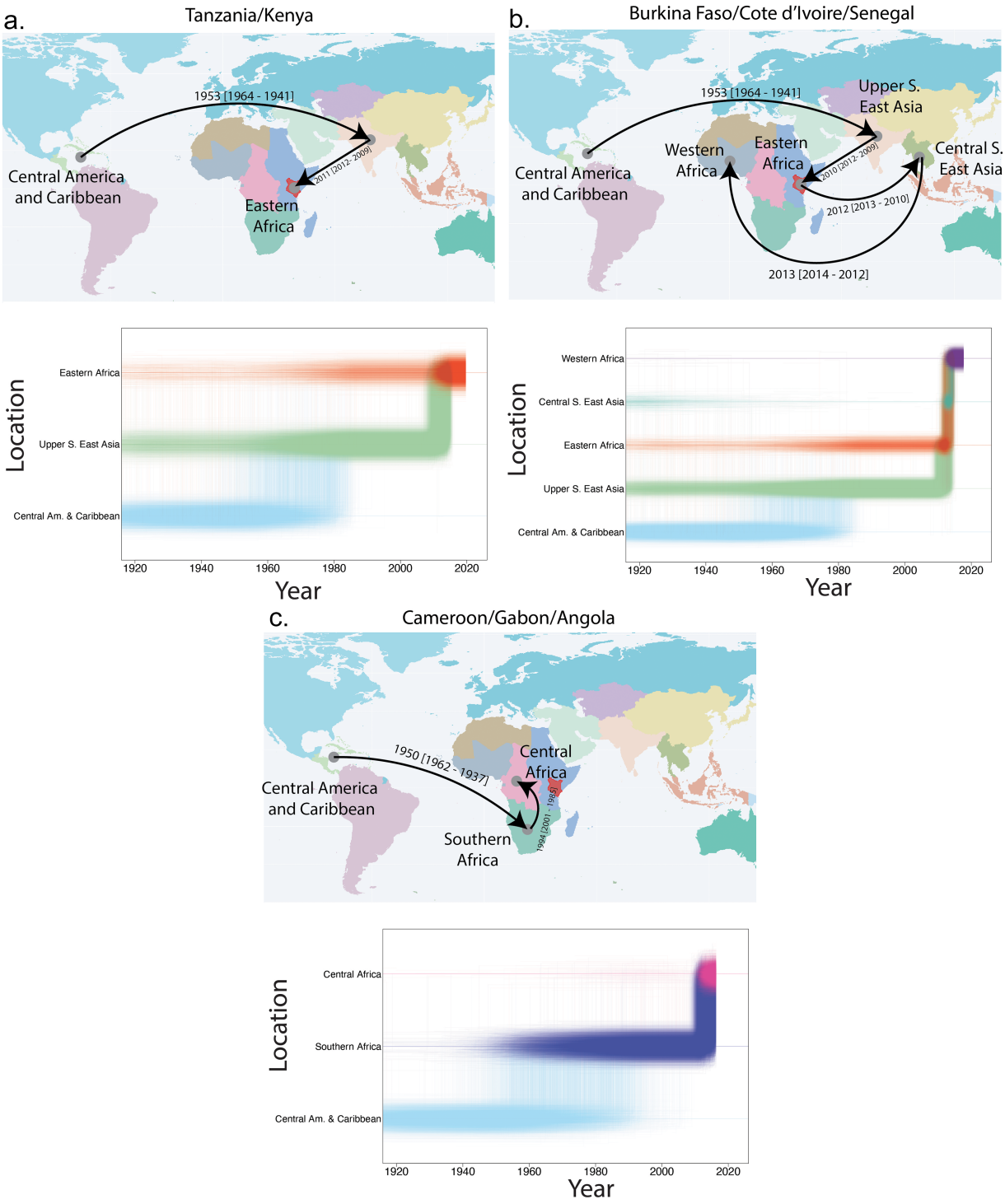

Figure 6. Geographic history of three DENV-1 clades. Bayesian MCC trees from phylogeographic analyses of 332 DENV-1 sequences were used to reconstruct geographic history of the three clades circulating in Africa. Each arrow in the map corresponds to a node on the MCC tree ancestral to the detected clade indicated in the plot, where a geographic transition between two locations occurred labeled with the inferred node date and 95% HPD interval. In order to represent the uncertainty in the timing of geographic transitions, Markov jumps were recorded along the history of the phylogeny and are shown in secondary plots for each clade. Vertical lines represent jumps or inferred transitions from the lower to the upper locations, while horizontal lines represent reward times or times during which clades are inferred to have remained in the same locations. The distribution of vertical lines represents uncertainty in geographic transition dates, with sparse lines indicating high uncertainty around a date and dense lines representing lower uncertainty around a date.

and Angola (Sessions et al. 2013, Hill et al. 2019). These outbreaks have been characterized using molecular surveillance although there were likely several other outbreaks that went undetected (Vu et al. 2017). In this analysis, we use a set of novel sequences collected from study sites in Kenya from 2019 to 2020, along with global sequence datasets to investigate the spread of DENV-1 and

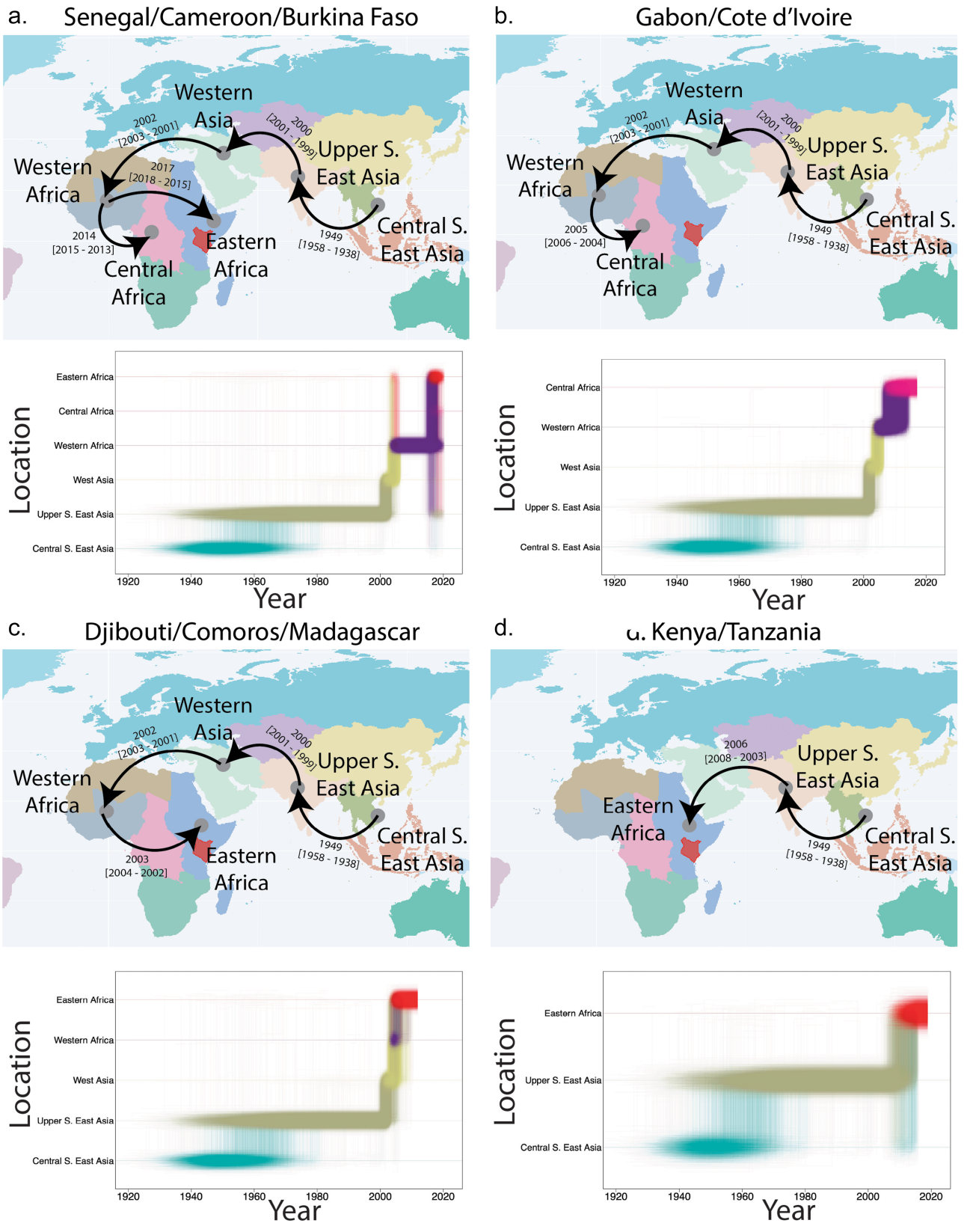

Figure 7. Geographic history of four DENV-3 clades. Bayesian MCC trees from phylogeographic analyses of 396 DENV-3 sequences were used to reconstruct geographic history of the four clades circulating in Africa. Each arrow in the map corresponds to a node on the MCC tree ancestral to the detected clade indicated in the plot, where a geographic transition between two locations occurred labeled with the inferred node date and 95% HPD interval. In order to represent the uncertainty in timing of geographic transitions, Markov jumps were recorded along the history of the phylogeny and are shown in secondary plots for each clade. Vertical lines represent jumps or inferred transitions from the lower to the upper locations, while horizontal lines represent reward times or times during which lineages are inferred to have remained in the same locations. The distribution of vertical lines represents uncertainty in geographic transition dates, with sparse lines indicating high uncertainty around a date and dense lines representing lower uncertainty around a date.

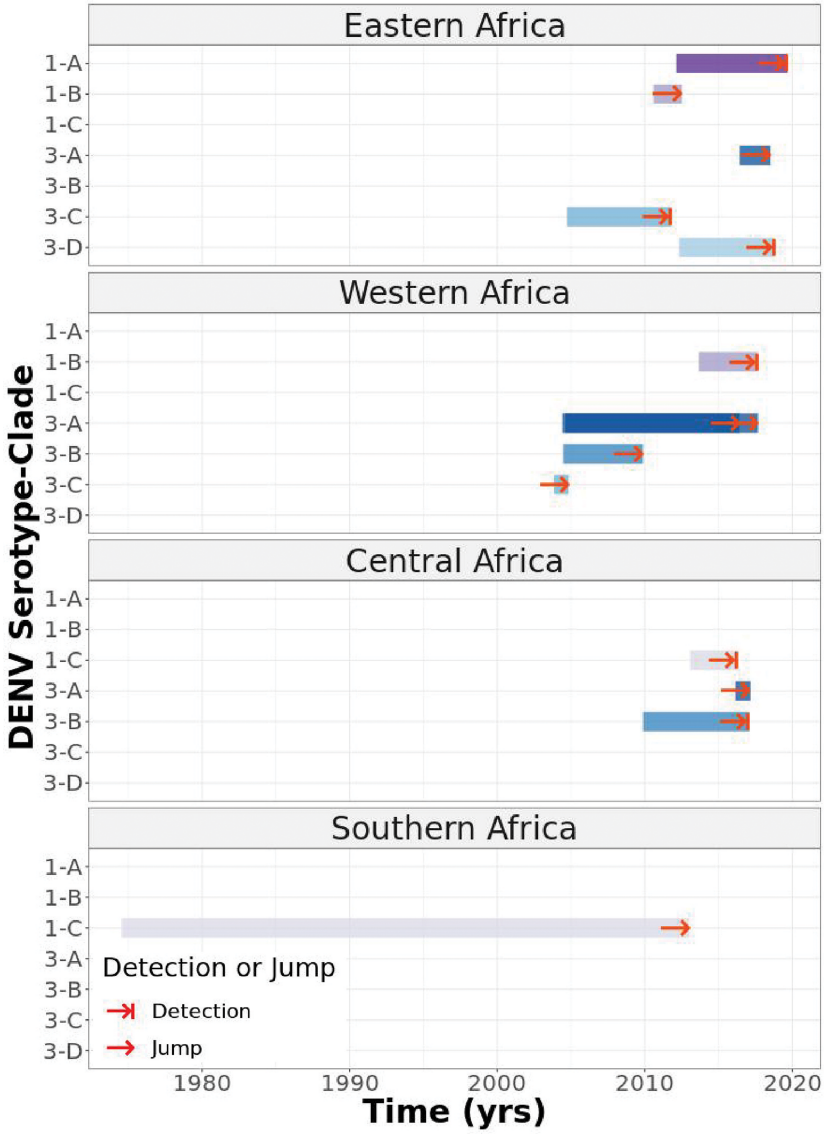

Figure 8. Identifying periods of cocirculation of DENV-1 and DENV-3 clades in Africa. The reward times from Markov jump analyses corresponding to study regions within Africa were summarized in order to investigate cocirculation and potential for hyperendemicity of DENV-1 and DENV-3 clades in Africa. Each bar in the plot denotes a summary of the time during which a clade circulated in the region. Circulation of a clade ends with either its detection during sampling (shown as an arrow with a bar: →|) or a "jump" from the original location to another location (shown as an arrow with no bar →). Start and end times for the circulation of each lineage were calculated using the median start and end times from the Markov jump history recorded along the phylogeny for each clade. Overlap of bars along the x-axis suggests cocirculation of two or more clades as seen in the Eastern and Western Africa regions.

DENV-3 in Kenya. We also analyze sequences collected from other locations in Africa to paint a complete picture of DENV expansion and circulation in the region.

Sequences collected from our study were closely associated with recent DENV-1 and DENV-3 outbreaks in Tanzania, indicating gene flow between the two locations. Our phylogenetic analysis resolved multiple clades of DENV-1 and DENV-3 circulating and causing outbreaks in East Africa as well as other regions (Figs 2 and 3). Most detected clades are recent importations (early to mid-2000s) from Asia (DENV-1 Clades a and b, DENV-3 Clades a, b, c, and d). The remaining DENV-1 clade is an older importation that seems to have circulated undetected since its putative introduction in the mid-1900s (DENV-1 Clade c). Our analysis shows that dengue circulation is characterized by a combination of local, within-region importations/exportations between adjacent or nearby countries and global importations/exportations involving Lower and Upper Southeast Asia and the Americas/Caribbean. Finally, our analyses reveal multiple geographic histories of DENV-1 and DENV-3 clades circulating in Africa, with distal origins in Lower Southeast Asia and the Americas (Figs 6 and 7).

The DENV-1 and DENV-3 isolates from this study were closely associated with samples collected from recent outbreaks in Tanzania, a 2017 DENV-3 outbreak and a 2019 DENV-1 outbreak (Fig. 4: DENV-1 Clade a and Fig. 5: DENV-3 Clade d) (Kelly et al. 2023). The four DENV-1 samples formed a monophyletic clade with an MRCA in 2019, suggesting a single importation from the Tanzania outbreak. The DENV-3 samples collected from this study and other studies in Kenya also resolved distinct clades for groups of samples collected from Kenya and Tanzania (Supplementary Fig. S5). This suggests that while there is gene flow between the two

Eastern African locations, their virus populations are not entirely panmictic and lineages circulate locally following initial regional importations. In addition to the Tanzania DENV-1 outbreak, the detected DENV-3 clade caused a reported 2019 outbreak in Kenya (Muthanje et al. 2022). These consecutive outbreaks in East Africa were caused by independent introductions of DENV-1 and DENV-3 from Upper Southeast Asia for DENV-1 and Western Asia for DENV-3 (Khan et al. 2020). Branch lengths suggest that these lineages circulated for ~ 5 years before detection during these outbreaks, and previous analyses of this outbreak support this finding (Muthanje et al. 2022) (Figs 2 and 4).

Due to the resolution of the phylogeographic analysis, our current results are unable to resolve the specific pathways of introduction of the clades circulating in East Africa. The clustering of sequences within the clades suggests that they were introduced in one of the two locations (Kenya or Tanzania) and spread to the other. Previous work has shown that Mombasa, an important regional port city in Kenya, is often the epicenter of dengue outbreaks and also plays a key role in DENV importations from Asia and exportations into other regions in Africa (Lutomiah et al. 2016, Lim et al. 2020). In order to identify the specific pathways of introductions from global dengue hotspots into regional nexus locations (such as Mombasa) and other locations using molecular epidemiology methods, more dense molecular surveillance at finer geographic resolutions is required. Identifying the specific pathways delineating global introductions from local spread would inform the design of effective interventions that would interrupt these pathways.

DENV-1 and DENV-3 circulation dynamics in other regions in Africa exhibit similar trends to those observed in East Africa. For example, many of the sequences isolated from West Africa form distinct geographic clades which are most closely related to clades from other nearby countries in the region and descended from a single importation from a region external to Africa (DENV-1 Clades a and b and DENV-3 Clades b and c). Many of these clades are recent introductions from Asia, with the exception of a single clade introduced from the Americas/Caribbean in the mid-1900s (DENV-1 Clade c). This pattern of independent importations into a region, and subsequent spread to nearby locations within that region, in contrast to multiple independent introductions into specific countries from global hotspots, is similar to observed DENV-2 dynamics in East Africa and indicates continuous importations of different lineages into Africa via different geographic pathways that are connected by nexus locations within regions. Our comparison of two different models of DENV-1 and DENV-3 circulation supports this, suggesting that DENV-1 and DENV-3 circulation routes may be governed by the same underlying process, centered around the same key locations. This is further supported by the rate matrixes from the primary models, which show that the same sets of locations contribute to DENV-1 and DENV-3 geographic spread, namely, East Asia and Lower Southeast Asia as well as South America. Targeted interventions at these key "nexus" or "source" locations may be a cost-effective way to reduce DENV expansion in Africa.

In addition to the clades highlighted in Figs 4 and 5, additional dengue sequences collected from Africa occur across the DENV-1 and DENV-3 phylogenetic trees, isolated at varying times since the late 1900s (Supplementary Figs S4 and S5). The earliest DENV-1 sequence included was collected in 1988 in Southern Africa, while the earliest DENV-3 sequence included was collected in 1985. These sequences as well as additional sequences scattered along the phylogenetic trees provide evidence of undetected DENV circulation in the region in the late 1900s, and early 2000s tied to circulation in Asia, based on the geographic history of these sequences. The lack of more widespread detection and characterization of early DENV circulation highlights the need for increased surveillance in Kenya and Africa more broadly.

Notably, despite only including the most well-represented clades of DENV-1 and DENV-3 in this analysis, our data suggest that there may be periods of cocirculation of DENV-1 and DENV-3 as early as the mid-2000s (Fig. 8). Visual inspection of the summarized Markov jump reward times shows that DENV-1 Clade a and DENV-3 Clade a cocirculated in East Africa before their near-simultaneous detection during surveillance of recent outbreaks, as previously reported (Kelly et al. 2023). Our analysis also highlights the cocirculation of DENV-1 Clade b and DENV-3 Clade a in West Africa between 2015 and 2020, with several isolates from both serotypes sampled from multiple countries in West Africa, including Senegal and Burkina Faso. Cocirculation of DENV-1 Clade c and DENV-3 Clade b is also observed in Central Africa, with sequences from both clades isolated from Gabon a few years apart (Fig. 4 Clade c and Fig. 5 Clade b). No cocirculation is observed in Southern Africa despite the inferred duration of the lineage's circulation in the region (Fig. 8). Early detection of both serotypes in the region (one DENV-3 sequence isolated from Mozambique in 1985 and three DENV-1 sequences isolated from Angola in 1988) suggests that there may have been periods of cocirculation in this region as well (Supplementary Figs S5 and S6).

Introduction and local cocirculation of multiple DENV serotypes and genotypes, leading to hyperendemicity, have important implications for disease control. While vector reduction is the primary form of disease control for DENV, vaccine interventions have also been proposed and tested in certain high-risk regions, and several vaccine candidates are currently under Phase III trials, with only two, Dengvaxia (CDY-TDV) and QDenga (TAK-003), approved for limited use in certain regions (Pintado Silva and Fernandez-Sesma 2023). Despite approval, vaccine efficacy has been shown to be complex, particularly in naïve populations or nonendemic settings (Flasche et al. 2016). Our data demonstrate that multiple DENV serotypes and lineages may circulate in populations undetected, complicating the determination of population immunity status for the purpose of dengue vaccination. As the range of DENV expands and more areas become suitable for transmission, silent circulation may become a barrier to accurately capturing population immune status, which is required for safe and effective vaccination campaigns.

Clade assignments within DENV genotypes in this analysis were used to delineate locally circulating clades within the African regions of interest, as no clear lineage designation system for DENV virus existed. Hill *et al.* have recently proposed a novel lineage assignment system for DENV that further assigns major and minor lineages within DENV genotypes (Hill et al. 2024). Using this new system, DENV-3 Clades a and b, identified in this analysis, are assigned Major lineage B and Minor lineage 2 (abbreviated 3III_B.2), while Clades c and d are assigned Major lineage B (abbreviated 3III_B). The majority of the DENV-1 Clade a samples are classified as Major lineage A and Minor lineage 2 (1III_A.2), with the four samples collected from our study classified as 1I_H, 1III_A, and 1I_B.

At the genotype level, the majority of samples fall into the same genotype in both the older and newly proposed classification systems, with a few exceptions, such as DENV-1 Clade c. This clade consists of samples collected from Central and Southern Africa between 2010 and 2021 and, according to the new classification, may fall into a new DENV-1 genotype,

Genotype VII. A complete list of major and minor lineage classifications according to this new proposed system is included in the Supplementary data (Supplementary Tables S4 and S5). In the absence of a widely adopted subgenotype classification system, multiple studies investigating DENV circulation in the region have resorted to applying different systems based on the study goals, as we have done here (Langat et al. 2020, Nyathi et al. 2024). A unified system of DENV classification at the sub-genotype levels, such as proposed by Hill *et al.*, will enable more uniform approaches to investigating DENV diversity at sub-genotype levels.

While this analysis presents a broad multiserotype assessment of DENV in Kenya and Africa, it has several limitations. Due to the low level of DENV surveillance in the region, while we are able to investigate contemporary patterns of circulation and spread, we are unable to compare these to earlier circulation patterns and determine how circulation dynamics may be shifting due to changing human, ecological, or evolutionary drivers. As anthropogenic environmental change continues, understanding how disease systems and their resulting dynamics shift over time will be vital for disease control. In these analyses, we are also only able to use a small number of sequences collected in our study and other studies in the region. As such, while we have made a contribution to the sequencing data available for these DENV serotypes, the total available sequences still pose a challenge to more granular analyses of any heterogeneous disease dynamics between the DENV serotypes within Kenya and Africa.

An additional limitation is that our analyses are only able to use data collected from recent significant outbreaks in large urban areas (Mombasa–Kenya, Ouagadougou–Burkina Faso, Thies–Senegal, etc.). Surveillance has shown that DENV is exported to and circulates in smaller, rural, and periurban locations (Masika et al. 2020). While DENV generally occurs in urban areas, rural locations may occur in close proximity to arboviral emergence and spillover zones. Human proximity with nonhuman primates and multihost vectors such as the forest dwelling, *Aedes albopictus* mosquito, found in West Africa, may enable rural locations to facilitate the emergence or spillback/spillover of DENV into/from sylvatic cycles (Longbottom et al. 2023). Without additional data and increased surveillance of rural and periurban locations, these hypotheses will be difficult to investigate.

In this analysis, we use dengue sequence data collected in 2020 from study sites in Kenya, along with sequences collected from other regions in Africa, to show that dengue circulation in East Africa and Africa more broadly is characterized by cocirculation of multiple serotypes and clades within a serotype. We show that most of these lineages were introduced recently from Asia via independent geographic pathways that focus on key locations in Africa although our resolution is not able to specifically identify these locations. Additional DENV surveillance is required to identify these key locations facilitating introductions and spread in the region, and interventions should be designed to interrupt DENV spread in these locations. Furthermore, the simultaneous contribution of multiple clades introduced at varying points to the observed dengue outbreaks suggests that dengue outbreak potential may be strongly influenced by factors other than virus lineage and corresponding evolutionary characteristics, reinforcing the importance of interrupting transmission links at the human and ecological level via interventions such as vector control.

Acknowledgements

The authors would like to thank all the study participants who contributed sample material to this study from the communities in Kisumu, Ukunda, Msambweni, and Chulaimbo (Kenya). They would also like to thank Dr M. K. Sahoo and Dr R. Verma for their input in refining the laboratory methods. They would also like to thank Dr Verity Hill for her feedback on the model comparison analysis.

Author contributions

S.N. conducted the laboratory work, bioinformatics, data analysis, and writing. I.M.R., K.S.W., and P.T. contributed to and supervised laboratory work and bioinformatic analysis and provided input on writing. F.M., B.N., D.M.V., and A.D.L. supervised recruitment, sample collection, and follow-up of study participants. J.O.M., P.A., and D.M.V. recruited study participants and collected samples and demographic data from study participants. E.A.M., S.B., J.R.A., and A.D.L. supervised the research project and contributed to study design, data analysis, writing and review of the manuscript.

Supplementary data

Supplementary data is available at *VEVOLU Journal* online.

Conflict of interest: None declared.

Funding

S.N. was funded by the Abbott Laboratories Graduate Student Fellowship and the Centre for Computational, Evolutionary and Human Genetics Fellowship at Stanford University. A.D.L. was funded for this study by an NIH grant (NIH Project Number: R01AI102918).

Data availability

The sequences generated as part of this study have been deposited to GenBank with accession numbers: OR765987–OR765990 and PQ412806 (https://www.ncbi.nlm.nih.gov/genbank/). All other sequences used in this analysis are available in the Supplementary data (Supplementary Tables S4 and S5). All .xml files generated as part of these analyses are available online (https://github.com/SindisoNyathi/doctoralrepo_dengue_phylo/tree/main/beast_xmls).

References

Bhatt S, Gething PW, Brady OJ *et al*. The global distribution and burden of dengue. *Nature* 2013;**496**:504–7.

Bosire C, Mutuku F, Ndenga B *et al*. A narrative review of dengue fever infection and epidemic activity in Kenya (2010 to 2020). *PAMJ One Health* 2023;**12**:10.

Brady OJ, Hay SI. The global expansion of dengue: how *Aedes aegypti* mosquitoes enabled the first pandemic arbovirus. *Annu Rev Entomol* 2020;**65**:191–208.

Chen RE, Smith BK, Errico JM *et al*. Implications of a highly divergent dengue virus strain for cross-neutralization, protection, and vaccine immunity. *Cell Host Microbe* 2021;**29**:1634–48.e5.

Cummings DAT, Irizarry RA, Huang NE *et al*. Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. *Nature* 2004;**427**:344–47.

Darriba D, Posada D, Kozlov AM *et al*. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol* 2020;**37**:291–4.

Ellis EM, Neatherlin JC, Delorey M *et al*. A household serosurvey to estimate the magnitude of a dengue outbreak in Mombasa, Kenya, 2013. *PLoS Negl Trop Dis* 2015;**9**:1–10.

Faye O, Ba Y, Faye O *et al*. Urban epidemic of dengue virus serotype 3 infection, Senegal, 2009. *Emerg Infect Dis* 2014;**20**:456.

Flasche S, Jit M, Rodríguez-Barraquer I *et al*. The long-term safety, public health impact, and cost-effectiveness of routine vaccination with a recombinant, live-attenuated dengue vaccine (Dengvaxia): a model comparison study. *PLoS Med* 2016;**13**:e1002181.

Fonseca V, Libin PJK, Theys K *et al*. A computational method for the identification of Dengue, Zika and Chikungunya virus species and genotypes. *PLoS Negl Trop Dis* 2019;**13**:e0007231.

Gainor EM, Harris E, LaBeaud AD. Uncovering the burden of dengue in Africa: considerations on magnitude, misdiagnosis, and ancestry. *Viruses* 2022;**14**:233.

Gaye A, Ndiaye T, Sy M *et al*. Genomic investigation of a dengue virus outbreak in Thiès, Senegal, in 2018. *Sci Rep* 2021;**11**:10321.

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**:1–19.

Halstead SB. Pathogenesis of dengue: challenges to molecular biology. *Science* 1988;**239**:476–81.

Harris PA, Taylor R, Thielke R *et al*. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Informat* 2009;**42**:377–81.

Hill SC, Vasconcelos JND, Granja BG *et al*. Early genomic detection of cosmopolitan genotype of dengue virus serotype 2, Angola, 2018. *Emerg Infect Dis* 2019;**25**:2017–20.

Hill V, Cleemput S, Pereira JS *et al*. A new lineage nomenclature to aid genomic surveillance of dengue virus. *PLoS Biol* 2024;**22**:e3002834.

Johnson B, Musoke S, Ocheng D *et al*. Dengue-2 virus in Kenya. *Lancet* 1982a;**2**:208–9.

Johnson B, Ocheng D, Gichogo A *et al*. Epidemic dengue fever caused by dengue type 2 virus in Kenya: preliminary results of human virological and serological studies. *East Afr Med J* 1982b;**59**:781–4.

Kelly ME, Msafiri F, Affara M *et al*. Molecular characterization and phylogenetic analysis of dengue fever viruses in three outbreaks in Tanzania between 2017 and 2019. *PLoS Negl Trop Dis* 2023;**17**:e0011289.

Khan NU, Danish L, Khan HU *et al*. Prevalence of dengue virus serotypes in the 2017 outbreak in Peshawar, KP, Pakistan. *J Clin Lab Analysis* 2020;**34**:e23371.

Kiener M, Shayegh N, Nyathi SV *et al*. Low rate of asymptomatic dengue infection detected in coastal Kenya using pooled polymerase chain reaction testing. *American JTrop Med Hyg* 2024;**110**:738–40.

Konongoi L, Ofula V, Nyunja A *et al*. Detection of dengue virus serotypes 1, 2 and 3 in selected regions of Kenya: 2011–2014. *Virol J* 2016;**13**:1–11.

Kozlov AM, Darriba D, Flouri T *et al*. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 2019;**35**:4453–5.

Labeaud AD, Peters CJ, King CH *et al*. Spectrum of Rift Valley fever virus transmission in Kenya: insights from three distinct regions. *American J Trop Med Hyg* 2007;**76**:795–800.

Langat SK, Eyase FL, Berry IM *et al*. Origin and evolution of dengue virus type 2 causing outbreaks in Kenya: evidence of circulation of two cosmopolitan genotype lineages. *Virus Evol* 2020;**6**:1–9.

Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 2014;**30**:3276–8.

Letizia AG, Pratt CB, Wiley MR *et al*. Retrospective genomic characterization of a 2017 dengue virus outbreak, Burkina Faso. *Emerg Infect Dis* 2022;**28**:1198.

Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 2010;**26**:589–95.

Li H, Handsaker B, Wysoker A *et al*. 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAM tools. *Bioinformatics* 2009;**25**:2078–9.

Lim JK, Matendechero SH, Alexander N *et al*. Clinical and epidemiologic characteristics associated with dengue fever in Mombasa, Kenya. *Inter J Infect Dis* 2020;**100**:207–15.

Longbottom J, Walekhwa AW, Mwingira V *et al*. Aedes albopictus invasion across Africa: the time is now for cross-country collaboration and control. *Lancet Glob Health* 2023;**11**:e623–8.

Lutomiah J, Barrera R, Makio A *et al*. Dengue outbreak in Mombasa City, Kenya, 2013–2014: entomologic investigations. *PLoS Negl Trop Dis* 2016;**10**:e0004981.

Masika MM, Korhonen EM, Smura T *et al*. Detection of dengue virus type 2 of Indian origin in acute febrile patients in rural Kenya. *PLoS Negl Trop Dis* 2020;**14**:e0008099.

Messina JP, Brady OJ, Scott TW *et al*. Global spread of dengue virus types: mapping the 70 year history. *Trends Microbiol* 2014;**22**:138–46.

Muthanje EM, Kimita G, Nyataya J *et al*. March 2019 dengue fever outbreak at the Kenyan south coast involving dengue virus serotype 3, genotypes III and V. *PLoS Global Public Health* 2022;**2**:e0000122.

Nyathi S, Rezende IM, Walter KS *et al*. Molecular epidemiology and evolutionary characteristics of dengue virus 2 in East Africa. *Nat Commun* 2024;**15**:7832.

Obonyo M, Fidhow A, Ofula V. Investigation of laboratory confirmed dengue outbreak in North-eastern Kenya, 2011. *PLoS ONE* 2018;**13**:e0198556.

Pintado Silva J, Fernandez-Sesma A. Challenges on the development of a dengue vaccine: a comprehensive review of the state of the art. *J Gen Virol* 2023;**104**:001831.

Pollett S, Gathii K, Figueroa K *et al*. The evolution of dengue-2 viruses in Malindi, Kenya and Greater East Africa: epidemiological and immunological implications. *Infect Genet Evol* 2021;**90**:104617.

Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 1998;**14**:817–8.

Rambaut A, Lam TT, Max Carvalho L *et al*. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* 2016;**2**:vew007.

Rico-Hesse R. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* 1990;**174**:479–93.

Sessions OM, Wilder-smith A, Meltzer E *et al*. Exploring the origin and potential for spread of the 2013 dengue outbreak in Luanda, Angola. *Global Health Action* 2013;**6**:21822.

Shah MM, Ndenga BA, Mutuku FM *et al*. High dengue burden and circulation of 4 virus serotypes among children with undifferentiated fever, Kenya, 2014–2017. *Emerg Infect Dis* 2020;**26**:2638–50.

Sutherland LJ, Cash AA, Huang YJS *et al*. Serologic evidence of arboviral infections among humans in Kenya. *Am J Trop Med Hyg* 2011;**85**:158–61.

Tariq A, Khan A, Mutuku F *et al*. Risk factors associated with dengue virus and chikungunya virus seropositivity and seroconversion among children in Kenya, a longitudinal study. *PLoS Negl Trop Dis* 2024;**18**:e0012616.

Ullrich KK. *MSA2dist: MSA2dist Calculates Pairwise Distances between All Sequences of a DNAStringSet or a AAStringSet Using a Custom Score*

Matrix and Conducts Codon Based Analysis (Version 1.8.0) [R]. 2024. https://gitlab.gwdg.de/mpievolbio-it/MSA2dist, accessed 15 Mar. 2023.

Vu DM, Krystosik AR, Ndenga BA *et al*. Detection of acute dengue virus infection, with and without concurrent malaria infection, in a cohort of febrile children in Kenya, 2014–2019, by clinicians or machine learning algorithms. *PLoS Global Public Health* 2023;**3**:e0001950.

Vu DM, Mutai N, Heath CJ *et al*. Unrecognized dengue virus infections in children, Western Kenya, 2014–2015. *Emerg Infect Dis* 2017;**23**:1915–7.

Waggoner JJ, Abeynayake J, Sahoo MK *et al*. Single-reaction, multiplex, real-time RT-PCR for the detection, quantitation, and serotyping of dengue viruses. *PLoS Negl Trop Dis* 2013;**7**: 1–9.

Waggoner JJ, Gresh L, Mohamed-Hadley A *et al*. Single-reaction multiplex reverse transcription PCR for detection of Zika, Chikungunya, and dengue viruses. *Emerg Infect Dis* 2016;**22**: 1295–7.

Zhang C, Mammen MP, Chinnawirotpisan P *et al*. Clade replacements in dengue virus serotypes 1 and 3 are associated with changing serotype prevalence. *J Virol* 2005;**79**:15123–30.

Virus Evolution, 2025, 11(1), veae116 , DOI: https://doi.org/10.1093/ve/veae116, Advance Access Publication 30 December 2024, Research Article