Virus Evolution, 2020, 6(1): veaa026

doi: 10.1093/ve/veaa026
Rapid Communication

Origin and evolution of dengue virus type 2 causing outbreaks in Kenya: Evidence of circulation of two cosmopolitan genotype lineages

Solomon K. Langat,[1,*,†,‡] Fredrick Lunyagi Eyase,[1,2,†] Irina Maljkovic Berry,[3] Albert Nyunja,[1] Wallace Bulimo,[1,4] Samuel Owaka,[1] Victor Ofula,[1] Samson Limbaso,[1,5] Joel Lutomiah,[1,5] Richard Jarman,[3] John Distelhorst,[1] and Rosemary C. Sang[1,5]

[1]Department of Emerging Infectious Diseases, United States Army Medical Research Directorate-Africa, Nairobi, Kenya, [2]Institute for Biotechnology Research, Jomo Kenyatta University of Agriculture and Technology, [3]Viral Diseases Branch, Walter Reed Army Institute of Research, [4]Department of Biochemistry, University of Nairobi and [5]Center for Virus Research, Kenya Medical Research Institute

*Correspondence author: E-mail: solomon.langat@usamru-k.org

†These authors contributed equally to this work.

‡https://orcid.org/0000-0002-9108-2261

Abstract

Dengue fever (DF) is an arboviral disease caused by dengue virus serotypes 1-4 (DENV 1-4). Globally, DF incidence and disease burden have increased in the recent past. Initially implicated in a 1982 outbreak, DENV-2 recently reemerged in Kenya causing outbreaks between 2011 and 2014 and more recently 2017–8. The origin and the evolutionary patterns that may explain the epidemiological expansion and increasing impact of DENV-2 in Kenya remain poorly understood. Using whole-genome sequencing, samples collected during the 2011–4 and 2017–8 dengue outbreaks were analyzed. Additional DENV-2 genomes were downloaded and pooled together with the fourteen genomes generated in this study. Bioinformatic methods were used to analyze phylogenetic relationships and evolutionary patterns of DENV-2 causing outbreaks in Kenya. The findings from this study have shown the first evidence of circulation of two different Cosmopolitan genotype lineages of DENV-2; Cosmopolitan-I (C-I) and Cosmopolitan-II (C-II), in Kenya. Our results put the origin location of C-I lineage in India in 2011, and C-II lineage in Burkina Faso between 1979 and 2013. C-I lineage was the most isolated during recent outbreaks, thus showing the contribution of this newly emerged strain to the increased DENV epidemics in the region. Our findings, backed by evidence of recent local epidemics that have been associated with C-I in Kenya and C-II in Burkina Faso, add to the growing evidence of expanding circulation and the impact of multiple strains of DENV in the region as well as globally. Thus, continued surveillance efforts on DENV activity and its evolutionary trends in the region, would contribute toward effective control and the current vaccine development efforts.

Key words: Kenya; dengue virus 2; evolution; origin; phylogenetics.

© The Author(s) 2020. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

1. Introduction

DF is an arboviral disease caused by dengue virus (DENV). DENV belongs to the family Flaviviridae, genus *Flavivirus*. It is a mosquito-borne pathogen that is traditionally transmitted by *Aedes aegypti* mosquito and to a lesser extent, *Aedes albopictus* (WHO 2014). In terms of disease burden and risk potential, DENV is one of the most important arboviral pathogens. One-third of the world's population (~2.5 billion people) is at risk of infection with DENV and it is estimated that approximately 390 million infections occur yearly in the tropical and subtropical regions (Bennett et al. 2003; Bhatt et al. 2013). DENV comprises four antigenically distinct serotypes, DENV 1-4. Infection with either of the DENV serotypes generally causes dengue fever (DF), a self-limiting febrile illness that often lasts between 2 and 10 days (Weaver and Vasilakis 2009). Some patients with DF can progress to develop one of two life-threatening conditions, dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS). DHF is characterized by hemorrhage and thrombocytopenia while DSS occurs due to excessive plasma leakage (Gubler 1994). Case fatality ratios due to DENV infections vary across populations, with an average of 5 percent being reported in some studies (Martin et al. 2016). Some of the factors contributing to these variations and the risk of development of life-threatening conditions include the genetic constitution of the virus, the genetic make-up of the individual as well as previous infection with a different serotype (Halstead 1999; WHO 1999). The spread and co-circulation of the serotypes have caused epidemics with increasing frequency and magnitude. Efforts to control the virus still depend on vector control strategies, given that the recent attempts to introduce a dengue vaccine have faced numerous challenges and have resulted in little success (Webster, Farrar, and Rowland-Jones 2009; Murray, Quam, and Wilder-Smith 2013; Thomas and Rothman 2015).

DENV has a single-stranded, positive sense nonsegmented RNA genome. Like many other RNA viruses, DENV is fast evolving due to the error-prone RNA-dependent RNA polymerase, leading to the formation of several genotypes within each DENV serotype over the years (Weaver and Vasilakis 2009; Waman et al. 2016). This evolution of the virus has been correlated with epidemic activity, and occasionally severity of the disease caused by the virus (Gubler 1998; Bennett et al. 2003; Bennett et al. 2006). DENV-2 has been divided into six different genotypes that include Asian I (AI), Asian II (AII), Cosmopolitan (C), American (AM), Asian/American (AA), and sylvatic (S) genotypes. These genotypes were suggested based on phylogeny of DENV-2 using envelope gene sequences (Twiddy et al. 2002). Recent advances in full-genome sequencing, however, have led to generation of complete DENV-2 genomes which has revealed further diversity within DENV-2 genotypes and thus suggestions for further groupings (Ali and Ali 2015; Waman et al. 2016). To ensure consistency throughout this study, the Cosmopolitan genotype has been divided into three lineages; Cosmopolitan-I (C-I), Cosmopolitan-II (C-II), and Cosmopolitan-III (C-III) (Ali and Ali 2015).

In Kenya, DF has been reported since World War II (McCarthy and Wilson 1948), with serological evidence recorded between 1966 and 1968 (Geser, Henderson, and Christensen 1970) suggesting that the virus had been circulating in coastal Kenya during that time. However, it was not until 1982 that the first dengue outbreak was documented in the Coastal towns of Malindi and Kilifi, which was subsequently determined to have been driven by DENV-2 (Johnson et al. 1982a,b). After this initial outbreak, no case of dengue disease was documented in this region for a period of almost two decades. In 1997, DENV-2 was isolated from a child with febrile illness in Kilifi, Kenya (Sang and Dunster 2001), suggesting possible continued active transmission of DENV in the region. This was further reinforced by the detection in 2004 of DENV IgG antibodies from human serum in Malindi (Mease et al. 2011). Between 2011 and 2014, an outbreak involving DENV-1, 2, and 3 was reported in Mandera, northeastern Kenya, and Mombasa, coastal Kenya (Konongoi et al. 2016; Obonyo, Fidhow, and Ofula 2018). In 2017, another outbreak involving DENV-2 was reported in Wajir and Mandera in North Eastern Kenya as well as Malindi, Kilifi, and Mombasa, on the Kenyan Coast (WHO 2017).

Considering the role of DENV-2 in the epidemiology of dengue disease in Kenya, we investigated the spatial origin and evolutionary inferences of the Kenyan DENV-2 lineages collected during the 2013–4 and 2017 outbreaks. The findings from this study extend our understanding of the origin of the two circulating lineages of DENV-2, C-I and C-II, in Kenya. Furthermore, the newly generated sequences add to the number of DENV-2 genomes currently available from Africa, where the number of genomes remains limited despite the numerous reported outbreaks.

2. Methods

2.1 Ethical consideration

The study was carried out on a protocol approved by the Walter Reed Army Institute of Research (#2189) and the Kenya Medical Research Institute's Institutional Review Boards (#3035), as an overarching protocol guiding investigation and reporting of arbovirus/hemorrhagic fever outbreaks in Kenya. Because blood samples were collected from an outbreak and no human data were collected, this was deemed a nonhuman research study and no consent was required. The protocol was approved for additional analysis of outbreak samples and for publication of results.

2.2 Samples

Samples sequenced in this study were collected during dengue outbreaks between 2013 and 2017. Serum samples were received at the Arbovirus and Viral Hemorrhagic Fever (VHF) laboratory, Kenya Medical Research Institute (KEMRI), and stored at −80°C.

2.3 RNA extraction and sequencing

Original specimen material stored at −80°C was aliquoted and used for viral RNA extraction. Viral RNA was extracted using QiaAmp Viral RNA Minikit (Qiagen, MD, USA), according to the manufacturer's instructions. Following extraction, Superscript III cDNA Synthesis Kit (Invitrogen, Carlsbad, CA, USA) was used for first-strand cDNA synthesis. Twelve primer pairs (Supplementary Table S1) were used to amplify 12 overlapping targets covering the entire genome of DENV-2. The amplicons were used to prepare libraries for next generation sequencing (NGS). In brief, the amplicons from each sample were pooled in equal volumes and cleaned using Agencourt Ampure XP beads (Beckman Coulter, Beverly, MA, USA) then quantified using Qubit 3.0 dsDNA HS Assay Kit (Thermo Fisher Scientific Inc., Wilmington, DE, USA). Libraries were then prepared using the Nextera XT DNA Sample Prep Kit (Illumina, San Diego, CA, USA), using 1 ng of DNA from the cleaned amplicons. The samples were uniquely barcoded using Nextera XT Index Kit (Illumina,

San Diego, CA, USA). Library normalization was done using the standard library normalization process according to the manufacturer's instructions. The normalized libraries were pooled in equal volumes and sequenced on an Illumina Miseq platform, using a 2 × 300 base paired-end reads. Raw sequence reads were inspected for quality using FASTQC (Andrews 2010). Prinseq-lite-0.20.4 (Schmieder and Edwards 2011) was used to filter the low-quality reads and *de novo* sequence assembly was performed with SPAdes v3.10 (Nurk et al. 2013). Reference-guided assembly was performed with Ngs_mapper pipeline (Cherokee Nation Technology Solutions 2015). The assembly followed default criteria set by the pipeline whereby the minimum base-calling quality of 25 and minimum mapping quality of 25 as well as a depth of coverage of 10 was used. Furthermore, every unambiguous nucleotide that was called had to have occurred in at least 80% of the assembled reads in the assembly. The sequences generated in this study were submitted to GenBank under accession numbers [MN577551–MN577564].

2.4 Sequence collation and screening of recombinants

Genomes from this study were combined with previously sequenced genomes from the Kenyan coast covering the period 2013–7 (Konongoi et al. 2016; Gathii et al. 2018; Kamau et al. 2019). In addition, DENV-2 sequences were downloaded from Virus Pathogen Resource (ViPR) and used for the initial phylogenetic analysis. These sequences belong to the different genotypes of DENV-2, excluding the sylvatic strains. For a more in-depth evolutionary analysis, we downloaded all the cosmopolitan genotype sequences from ViPR, containing the date of isolation and country of origin. Sequences that did not have this information were manually assigned through literature search and those that could not be assigned were excluded. To normalize the number of sequences from the individual countries, we used CD-HIT (Fu et al. 2012) to cluster the sequences from Countries that had a disproportionately high number of sequences, using a similarity threshold of 0.98–0.99 depending on the number of clusters from each country. We randomly picked one sequence from each cluster and used them for subsequent analyses.

Alignments were obtained using MAFFT v7.428 (Katoh and Standley 2013), Jalview v2 (Waterhouse et al. 2009) was used to manually edit the alignments. The sequences were then trimmed to only remain with the open reading frames of DENV-2, with further trimming being performed to remove sites that had long stretches of missing bases, leading to a final alignment of 7,040 nt. Recombination detection was performed on each of these alignments in the RDP4 software suite (Martin et al. 2015) using RDP, GENECOV, and MAXCHI methods. RECSCAN and SISCAN methods were used for confirmatory recombination scans. Significant recombination was considered when it was detected by all the three methods and at least one confirmatory method. We removed the sequences that were detected to be recombinants based on these criteria. Each of the alignments was further confirmed by using the Genetic Algorithm for Recombination Detection (GARD) (Kosakovsky Pond et al. 2006) in datamonkey webserver. No evidence of recombination was found in the final alignments.

2.5 Phylogenetic analysis

Maximum Likelihood (ML) and Bayesian phylogenies were inferred using RAxML (Stamatakis 2014) and MrBayes V3.2.6 (Ronquist et al. 2012), respectively. The best-fit substitution model for our datasets was determined by both the AIC and BIC tests in jModelTest2 (Darriba et al. 2012) as the general-time reversible model with gamma distribution (GTR+ G). This model was applied in both analyses. RAxML was performed for 2000 bootstrap replications and MrBayes implemented the Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) method, with three hot chains and one cold chain, and the analysis was run for 5 million generations with 25% burn-in. The analysis was performed in duplicate and the convergence of runs was determined by the average standard deviation in split frequencies of less than 0.01. Because majority of the additional DENV-2 genomes from the 2013 to 2014 outbreak period were only sequenced on their capsid/premembrane (C/PrM) gene region, we also inferred the phylogenetic tree for the partial C/PrM region. All the additional C/PrM sequences from Kenya as well as other closely related C-I and C-II sequences were downloaded and combined with the BEAST dataset (C/PrM region only). ML trees were then inferred, using the GTR+G, as determined by jModelTest2.

2.6 Molecular clock analysis and phylogeography

Time-calibrated phylogeny and phylogeography analyses were performed using the Cosmopolitan genotype sequence dataset only. Prior to performing this analysis, TempEst v1.5.3 (Rambaut et al. 2016) was used to assess the clock-likeness of our data by estimating the root-to-tip regression. We found a strong signal (correlation coefficient = 0.7923), suggesting the dataset was appropriate for the estimation of temporal parameters. The analysis was then carried out using the Bayesian Markov Chain Monte Carlo (MCMC) method, implemented in BEAST V1.10.4 (Suchard et al. 2018). For a more tractable postanalysis interpretation and reduced computational complexity, we collapsed the country of origin of the sequences to five distinct regions as their trait location; Africa, SE Asia, South Asia, East Asia, and Oceania. The Kenyan Sequences were assigned 'Kenya' as their trait location. We applied a relaxed molecular clock (uncorrelated lognormal) and a nonparametric Bayesian Skyline model, with ten groups, to estimate past population dynamics. The best-fit substitution model of GTR+G, as suggested by jModelTest2 (Darriba et al. 2012), was applied in this analysis. An asymmetric trait substitution model was used and reconstruction of ancestral states at all nodes was activated. Two independent runs of 250 million each, sampling after every 25,000 steps, were performed and the runs were combined after removal of 10 percent of the trees as burn-in. The convergence of runs was checked with Tracer v1.7.1 (Rambaut et al. 2018) to ensure that adequate mixing was achieved and that all parameters had reached an adequate effective sample size (ESS > 200). The maximum-clade credibility tree was summarized with TreeAnnotator and visualized in FigTree v1.4.4.

3. Results

Overlapping PCR amplicon targets for fourteen DENV-2 isolates were obtained and subjected to NGS. We generated consensus sequences for all fourteen isolates which had varying lengths of between 8,424 and 10,292 nt, with an average length of 9,562 nt. Thirteen of the isolates sequenced in this study originated from Mombasa and one originated from Wajir, northeastern Kenya. The Wajir isolate was collected in June of 2014, during the northeastern Kenya dengue outbreak (Konongoi et al. 2016; Obonyo, Fidhow, and Ofula 2018). We downloaded a further fourteen full-genome sequences from the Kenyan coast, which

were then combined with the sequences generated in this study for subsequent analysis.

3.1 C-I and C-II lineages of the cosmopolitan genotype were circulating in Kenya

Both ML and Bayesian phylogenies were concordant. They showed that 93% (26 of the 28) of the isolates from Kenya fell within a single and well-supported monophyletic clade (Bootstrap, BS = 100; Posterior probability, PP = 1) (Fig. 1A and Supplementary Fig. S1). This clade clustered under the C-I lineage of the Cosmopolitan genotype and was closely related to isolates from South Asia; more specifically India, Pakistan, and Sri Lanka. Our phylogenetic analysis also revealed the circulation of a comparatively divergent strain of DENV-2. This strain was represented by one isolate from Mombasa and one isolate from Wajir, which was isolated during the 2013–4 dengue outbreak in the northeastern part of Kenya (BS = 100, PP = 1) (Fig. 1A and Supplementary Fig. S1). The Mombasa/Wajir genomes clustered under the C-II lineage of the Cosmopolitan genotype. They were closely related to genomes sampled in Burkina Faso from 1983 to 2016 and those from Indonesia sampled in 1975–6 as well as a genome from one case originating from Niger (Supplementary Fig. S1). The long branch leading to this Kenyan strain suggests a lack of data that would provide a clearer picture of the geographic distribution of this particular strain.

ML phylogeny for the partial C/PrM gene, including additional sequences from Kenya, recovered one additional sequence from Mombasa that belonged to the C-II lineage (Fig. 1B). The other additional five sequences, all from the Kenyan Coast, fell under the C-I lineage (Fig. 1B). This finding agrees with our observation that C-I lineage was the most predominant during the 2013–4 outbreak and reinforces the evidence of the existence of C-II lineage in Mombasa during this outbreak period. There was, however, no additional C-II lineage sequences from any other additional country, pointing to the possibility of this strain currently being present only in the countries identified before; Burkina Faso, Niger, and Kenya.

3.2 Strains causing DENV-2 outbreaks in Kenya were introduced independently

The molecular clock analyses indicate that the most recent common ancestor (MRCA) of the coastal Kenya outbreak strain existed in 2011 (2011.57; 95% HPD: 2010.83, 2012.24) (Fig. 2), making this the most probable latest time point of introduction of this particular strain to the Kenyan coast. We estimated the evolutionary rates of this particular strain under the uncorrelated lognormal clock (UCLN) and our findings showed that the mean nucleotide substitution rates of this strain was 1.6631E−3 substitutions/site/year (95% HPD: 6.0638E−4, 3.1978E−3), which was not significantly different from the nucleotide substitution rates of the entire cosmopolitan genotype, 8.403E−4 substitutions/site/year (95% HPD: 7.4527E−4, 9.3025E−4). These findings were consistent with previous evolutionary rate estimates for DENV2, 8.94E−4 substitutions/site/year (95% HPD: 7.39E−4, 1.04E−3 (Wei and Li 2017). Discrete state reconstruction puts, with high confidence, the origin location of the ancestral strain in South Asia and most probably India, where most of the closely related isolates came from (PP = 0.9).

The MRCA of the Wajir/Mombasa strain existed in Kenya in early 2013 (2013.31; 95% HPD: 2013.07, 2013.38) (Fig. 2) with the discrete state reconstruction putting the origin location of this strain's MRCA in Africa (PP = 0.7). This strain shared an ancestor with genomes from Burkina Faso, which were isolated since 1983 and the ancestor existed in late 1979 (1979.82; 95% HPD: 1978.22, 1981.39). This strain was, therefore, introduced to Kenya between September 1979 and March 2013.

4. Discussion

DENV-2 has been the dominant serotype reported during dengue outbreaks in Kenya since the 1980s (Johnson et al. 1982a,b; Konongoi et al. 2016; Lutomiah et al. 2016; Gathii et al. 2018). The outbreak that occurred in 1982 was considered to be extensive and although little information is available as to the extent of this outbreak, one group tested forty-three acutely ill patients and obtained seven virus isolates (Johnson et al. 1982a,b). Furthermore, the antibody prevalence rate during this outbreak period rose from 7.6 in June, 1982 to 52 percent in October the same year (WHO 1993). The next set of outbreaks was reported between 2011 and 2014 and more recently between 2017 and 2018. The 2011–4 outbreak involved DENV-1, 2, and 3 (WHO 2011, Konongoi et al. 2016; Lutomiah et al. 2016) and the 2017 outbreak involved DENV-2 (WHO 2017; Gathii et al. 2018). These two outbreaks caused high public health burden, involving thousands of cases and multiple fatalities (ProMED-mail 2011; Njeru 2013; The New Humanitarian 2013; Ellis et al. 2015; WHO 2017). In this study, we investigated the evolutionary origins of DENV-2 that caused these recent outbreaks in Kenya by analyzing DENV-2 sequence data obtained during the outbreak in coastal Kenya and Wajir between 2013 and 2017.

The most significant finding from this study is the revelation of the presence of two different lineages of DENV-2 Cosmopolitan genotype, C-I and C-II, in Kenya. It is important to note that after the initial emergence of the Cosmopolitan genotype, there was early separation of C-I and C-II/C-III lineages, leading to emergence of the two clearly defined lineages. Our findings show that this early separation occurred by April 1958 (95% HPD: 1950.46, 1963.89), with the C-II and C-III lineages separating by January 1961 (95% HPD: 1956.85, 1964.63) (Fig. 2). It, therefore, follows that the two lineages circulating in Kenya had separated by April, 1958. Our phylogenetic analysis show that the majority of the isolates from the coastal Kenya outbreak (Coastal Kenya strain) fell within the C-I lineage and the isolate from Wajir, together with two isolates from Mombasa (Wajir/Mombasa strain), fell within the C-II lineage. This, therefore, points to a more significant impact of the C-I lineage during the recent DENV-2 outbreaks in Kenya as compared with the C-II lineage.

The C-I lineage contains DENV-2 genomes predominantly from South Asia. There are also sporadic isolations of this lineage in East Asia, particularly South Korea and China. Sequences isolated recently in the islands of Reunion and Seychelles also fell within this lineage. However, the strains that caused this Indian Ocean islands (Seychelles and Reunion) epidemic (Pascalis et al. 2019) do not appear to have any link to the coastal Kenya outbreak strain. Our analysis show that the C-I lineage was introduced into the Kenyan coast between March 2010 and June 2011 and it led to successful local transmission of DENV-2, as evidenced during the outbreak in Mombasa in 2013–4 (Konongoi et al. 2016; Lutomiah et al. 2016). This outbreak was responsible for at least 153 confirmed cases and 1 fatality (Njeru 2013; Ellis et al. 2015). The C-II lineage, on the other hand, clustered together with DENV-2 genomes from Burkina Faso and Indonesia as well as one genome from Niger. This lineage appears to have been introduced into Kenya between September 1979 and March 2013. It was isolated during both the

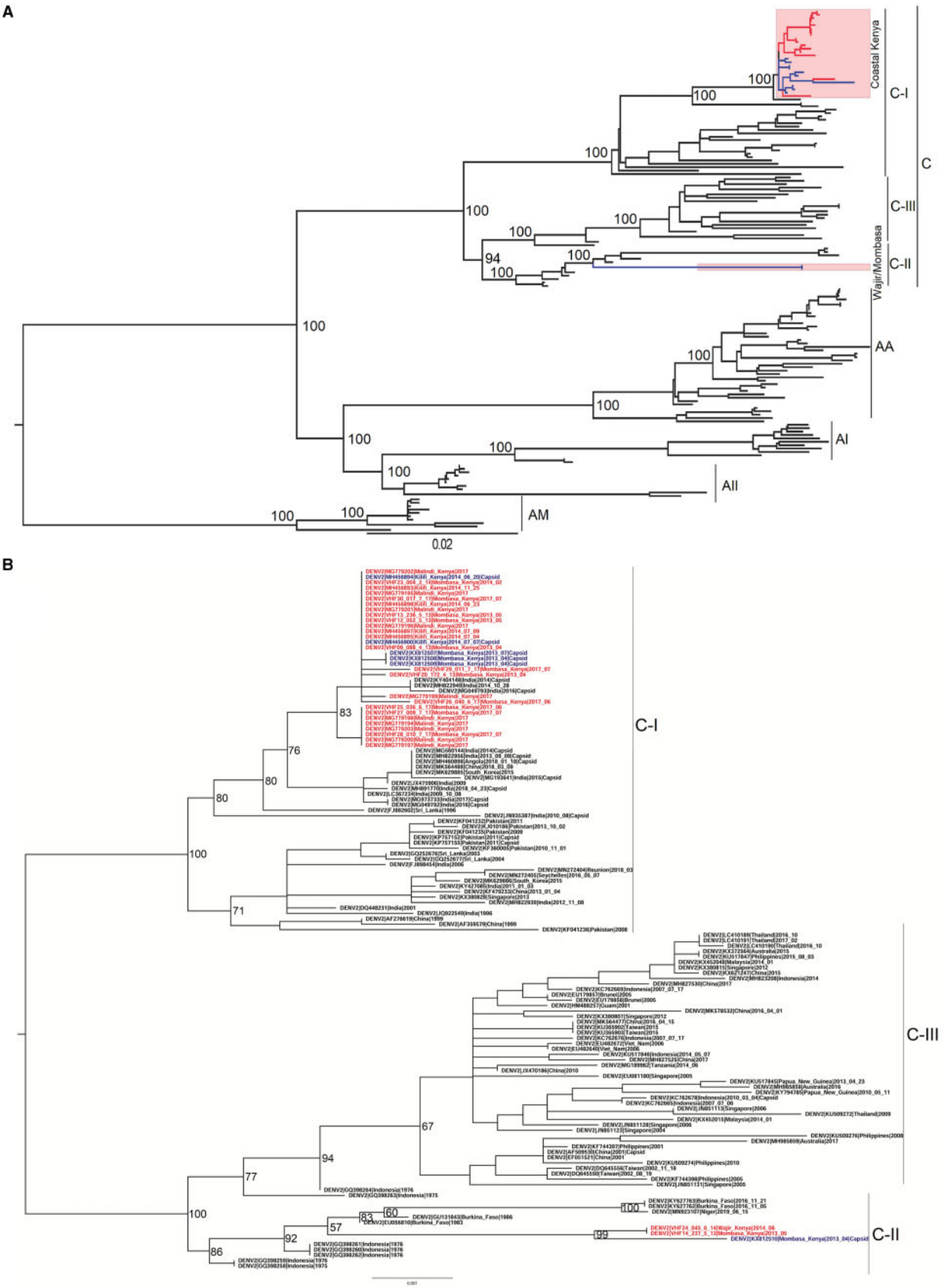

Figure 1. (A) Mid-point rooted ML phylogenetic Tree of 154 genome sequences isolated globally. Sequences from Kenya are highlighted with red tips representing sequences from 2017 outbreak and blue tips representing the 2013/4 outbreak sequences. The annotations shown are represented by; C, Cosmopolitan genotype; C-I, Cosmopolitan-I lineage; C-II, Cosmopolitan-II lineage; C-III, Cosmopolitan-III lineage; AA, Asian-American genotype; AI, Asian I genotype; AII, Asian II genotype; AM, American genotype. (B) ML phylogenetic tree generated using 129 C/PrM gene sequences of the Cosmopolitan genotype only. The additional Kenyan sequences that were included in this dataset are colored with blue tip-labels.

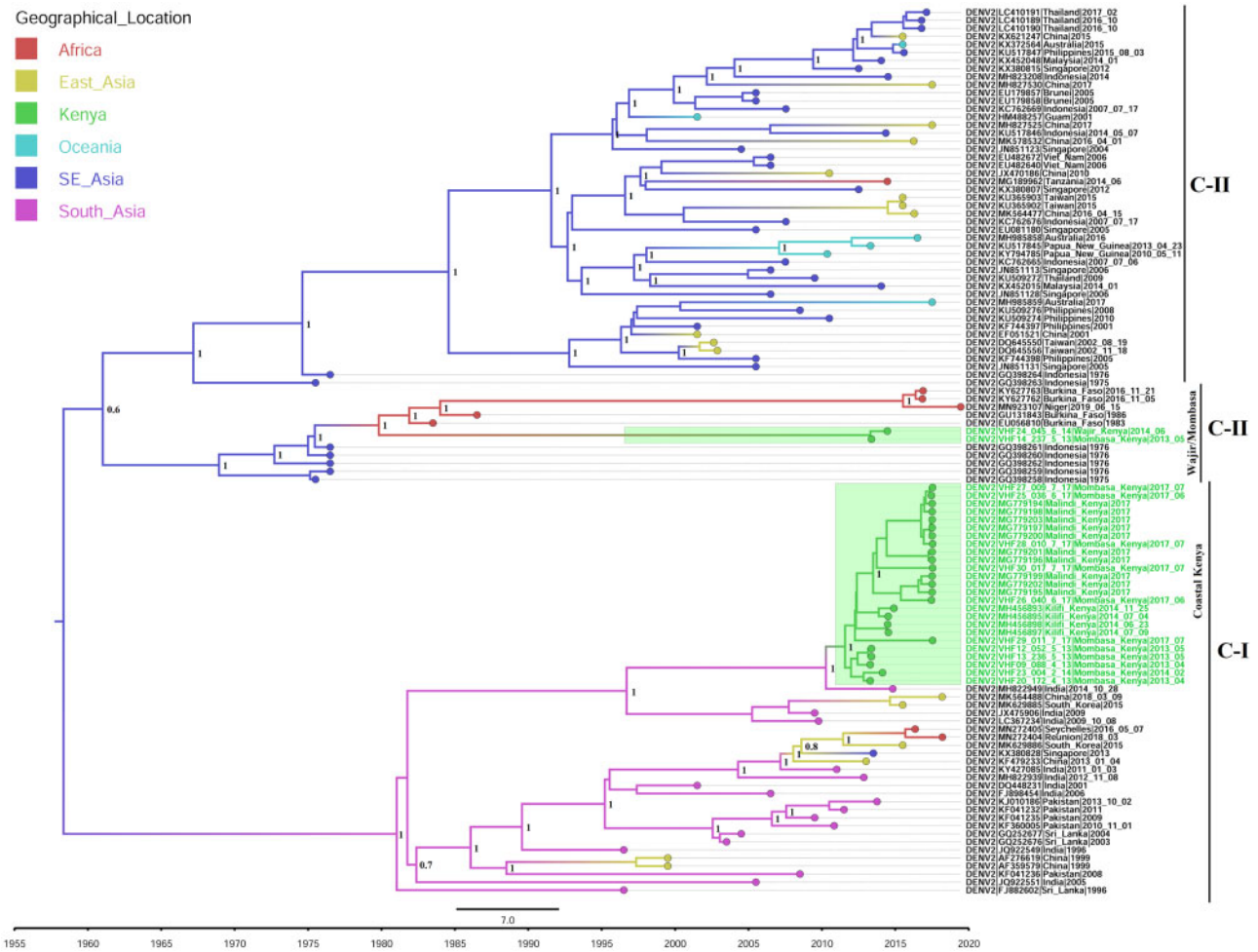

Figure 2. Time-calibrated phylogeny based on 110 cosmopolitan genotype sequences from different countries. The Kenyan sequences are colored in green tip-labels and the tree branches are colored based on the inferred geographical location of origin.

Coastal and northeastern Kenya outbreaks in 2013–4. The outbreak in northeastern Kenya in 2011 and in 2013 was responsible for at least 5,300 documented cases (with at least thirteen fatalities) (ProMED-mail 2011; The New Humanitarian 2013). Only one genome from Wajir was available in this study and, therefore, we cannot make any conclusion on the lineage or the DENV serotype that was mostly involved in that particular outbreak. Nonetheless, the Kenyan dengue outbreak between 2011 and 2014 involved DENV-1, 2, and 3 and the co-occurrence of these three serotypes may have played a major role in the extended nature of this outbreak as well as the increased number of cases and fatalities reported (Dhanoa et al. 2016). However, the contribution of DENV-2 was also quite significant considering that it was the most commonly isolated serotype during 2013 and 2014 (Konongoi et al. 2016).

DENV-2 further reemerged in both Mombasa and Wajir counties in 2017, after a period of more than 2 years. This outbreak also involved large number of cases, with 82 cases being reported in Wajir and a much larger number of at least 1,117 cases (with one fatality) being reported in Mombasa county as of July 2017 (WHO 2017). Our analysis of all the sequences from Mombasa and Kilifi Counties which were obtained during the outbreak in 2017 showed that all the isolates were of the C-I lineage. Put together, these findings show the establishment and expansion of the C-I lineage and its increased impact in regard to DENV outbreaks in Kenya.

Our analysis of the probable geographic origins of the currently circulating DENV-2 strains showed South Asia as the epicenter of C-I outbreaks, with evidence of seeding to other countries in East Asia, South East Asia, and much more recently in Africa (Kenya, Seychelles, and Reunion). The introduction of this strain into Kenya was estimated to have occurred from this particular region, and most probably from India where most of the closely related strains came from. This may be due to the interconnection between Mombasa and various countries in South Asia, particularly India (Bita 2015). The coastal city of Mombasa is East Africa's main shipping port and an international tourist destination; thus, it may serve as the entry point of DENVs and mosquito vectors through tourism and intercontinental trade. Indeed, a previous study on Chikungunya has supported this hypothesis whereby a strain of Chikungunya that caused an outbreak in northeastern Kenya in 2016 was found to have been introduced into Kenya from India (Maljkovic Berry et al. 2019). At the same time, seeding of Chikungunya to the Indian Ocean Islands and the Indian mainland was shown to have happened from this location (Volk et al. 2010), where it triggered one of the worst CHIKV epidemics between 2004 and 2007 (Borgherini et al. 2007; Chretien et al. 2007;

Kariuki Njenga et al. 2008). Our results, therefore, support this previous finding and can be explained by the fact that dengue and Chikungunya are transmitted by the same mosquito vector species. This intercontinental network between the Coastal city of Mombasa and countries in South Asia, therefore, provides a nexus for transmission of dengue between Mombasa and South Asia and may have been the channel through which C-I lineage was introduced into Kenya. The C-II lineage strain appears to have been introduced into Kenya possibly from Burkina Faso. However, the long branch as well as the estimated time of introduction between September 1979 and March 2013 points to missing data. Likewise, this lack of genomic data makes it impossible to estimate with certainty the geographical distribution of this strain over time. It is noteworthy to point out that the Wajir/Mombasa strain diverged from Burkina Faso strains in September 1979 (95% HPD: 1978.22, 1981.39). Considering the first documented DENV-2 outbreak at the Kenyan coast was in 1982 (Johnson et al. 1982a,b), it is plausible that the ancestral strains of the Wajir/Mombasa strain were responsible for DENV-2 outbreak during that period (Johnson et al. 1982a). However, more sequence data from previous Kenyan dengue outbreaks, would confirm whether this strain was responsible for the outbreak experienced in 1982 and later periods. The phylogenetic analysis showed the C-II lineage to have little representation with regard to the number of available sequences. Unlike the C-I lineage which has been implicated in numerous outbreaks across Asia and Africa, the only recent documented epidemic driven by C-II lineage is the Burkina Faso outbreak of 2016 (WHO 2016; Baronti et al. 2017). Prior to this outbreak, C-II had only been isolated in Indonesia in 1975–6 and in Burkina Faso in 1983. The lack of data for this particular lineage could be due to lack of surveillance in areas where it is most common, possibly in Africa and/or Asia. Alternatively, this could also be explained by the intrinsic factors that are particular to C-II lineage of DENV-2. Nevertheless, we cannot explain with certainty the reasons why this lineage is rarely isolated compared with the other Cosmopolitan genotype lineages. It is, however, worth noting that the recent outbreaks reported in Burkina Faso in 2016 were driven by this lineage (C-II) where it was associated with at least 1,061 cases with fifteen fatalities (WHO 2016; Baronti et al. 2017). This epidemic continued into 2017, where DENV-1, 2, and 3 were identified during this outbreak period (ReliefWeb 2017; WHO 2017). Considering DENV-2 was among the serotypes identified, it is possible that the C-II lineage was also responsible for the DENV-2 cases during this Burkina Faso 2017 outbreak.

5. Conclusion

This study has shown that recent epidemics associated with DENV-2 in Kenya were predominantly driven by the newly introduced C-I lineage, which was introduced to Kenya between March 2010 and June 2011 from South Asia, and which subsequently spread and contributed to the DENV-2 outbreak in 2013/4 and 2017. The intercontinental link of the spread, therefore, shows that both local and international strategies should be considered when designing interventions to curb DENV-2 transmission in Kenya and the region. We have also described a comparatively divergent strain of DENV-2, C-II, which may have been circulating in this region since the 1980s. Even though this strain was least isolated during the outbreak in 2013–4, its role in the 2016 Burkina Faso outbreak points to the potential for this strain to cause local epidemics and thus continued surveillance efforts on DENV activity are important for effective control.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Funding

This work was supported by the Armed Forces Health Surveillance Centre (AFHSC), Division of Global Emerging Infections Surveillance and Response System (GEIS) and FY2019 ProMIS ID P0136_19_KY.

Disclaimer

The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or reflecting the views of United States Army Medical Research Directorate-Africa, Kenya Medical Research Institute, Department of the Army, Department of Defense or the US Government. The investigators have adhered to the policies for the protection of human subjects as prescribed in AR 70-25.

Conflict of interest: None declared.

References

Ali, A., and Ali, I. (2015) 'The Complete Genome Phylogeny of Geographically Distinct Dengue Virus Serotype 2 Isolates (1944-2013) Supports Further Groupings within the Cosmopolitan Genotype', *PLoS One*, 10: e0138900.

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data* <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> accessed 6 Mar 2020.

Baronti, C. et al. (2017) 'Complete Coding Sequences of Two Dengue Virus Type 2 Strains Isolated from an Outbreak in Burkina Faso in 2016', *Genome Announcements*, 5: e00209–17.

Bennett, S. N. et al. (2003) 'Selection-Driven Evolution of Emergent Dengue Virus', *Molecular Biology and Evolution*, 20: 1650–8.

—— et al. (2006) 'Molecular Evolution of Dengue 2 Virus in Puerto Rico: Positive Selection in the Viral Envelope Accompanies Clade Reintroduction", *The Journal of General Virology*, 87: 885–93.

Bhatt, S. et al. (2013) 'The Global Distribution and Burden of Dengue', *Nature*, 496: 504–7.

Bita, C. (2015) 'Historical Period Stone Anchors from Mombasa, Kenya: Evidence of Overseas Maritime Trade Contacts with Asia and Middle East', *International Journal of Environment and Geoinformatics*, 2: 15–26.

Borgherini, G. et al. (2007) 'Outbreak of Chikungunya on Reunion Island: Early Clinical and Laboratory Features in 157 Adult Patients', *Clinical Infectious Diseases*, 44: 1401–7.

Cherokee Nation Technology Solutions. (2015). 'Ngs_mapper: Genome Mapping Pipeline' <https://github.com/VDBWRAIR/ngs_mapper> accessed 20 Jan 2020.

Chretien, J.-P. et al. (2007) 'Drought-Associated Chikungunya Emergence along Coastal East Africa', *The American Journal of Tropical Medicine and Hygiene*, 76: 405–7.

Darriba, D. et al. (2012) 'jModelTest 2: More Models, New Heuristics and Parallel Computing', *Nature Methods*, 9: 772.

Dhanoa, A. et al. (2016) 'Impact of Dengue Virus (DENV) co-Infection on Clinical Manifestations, Disease Severity and Laboratory Parameters', *BMC Infectious Diseases*, 16: 406.

Ellis, E. M. et al. (2015) 'A Household Serosurvey to Estimate the Magnitude of a Dengue Outbreak in Mombasa', *PLoS Neglected Tropical Diseases*, 9: e0003733.

Fu, L. et al. (2012) 'CD-HIT: Accelerated for Clustering the Next-Generation Sequencing Data', *Bioinformatics*, 28: 3150–3152.

Gathii, K. et al. (2018) 'Complete Coding Sequences of Dengue Virus Type 2 Strains from Febrile Patients Seen in Malindi District Hospital, Kenya, during the 2017 Dengue Fever Outbreak', *Genome Announcements*, 6: e00076–18.

Geser, A., Henderson, B. E., and Christensen, S. (1970) 'A Multipurpose Serological Survey in Kenya. 2. Results of Arbovirus Serological Tests', *Bulletin of the World Health Organization*, 43: 539–52.

Gubler, D. J. (1994) 'Perspectives on the Prevention and Control of Dengue Hemorrhagic Fever', *Gaoxiong yi Xue ke Xue za Zhi = the Kaohsiung Journal of Medical Sciences*, 10: S15–8.

—— (1998) 'Dengue and Dengue Hemorrhagic Fever', *Clinical Microbiology Reviews*, 11: 480–496.

Halstead, S. B. (1999) 'Is There an Inapparent Dengue Explosion?', *The Lancet*, 353: 1100–1101.

Johnson, B. K. et al. (1982a) 'Dengue-2 Virus in KENYA', *The Lancet*, 320: 208–9.

—— et al. (1982b) 'Epidemic Dengue Fever Caused by Dengue Type 2 Virus in Kenya: Preliminary Results of Human Virological and Serological Studies', *East African Medical Journal*, 59: 781–4.

Kamau, E. et al. (2019) 'Complete Genome Sequences of Dengue Virus Type 2 Strains from Kilifi, Kenya', *Microbiology Resource Announcements*, 8: e01566–18.

Kariuki Njenga, M. et al. (2008) 'Tracking Epidemic Chikungunya Virus into the Indian Ocean from East Africa', *Journal of General Virology*, 89: 2754–60.

Katoh, K., and Standley, D. M. (2013) 'MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability', *Molecular Biology and Evolution*, 30: 772–80.

Konongoi, L. et al. (2016) 'Detection of Dengue Virus Serotypes 1, 2 and 3 in Selected Regions of Kenya', *Virology Journal*, 13: 182.

Kosakovsky Pond, S. L. et al. (2006) 'GARD: A Genetic Algorithm for Recombination Detection', *Bioinformatics*, 22: 3096–8.

Lutomiah, J. et al. (2016) 'Dengue Outbreak in Mombasa City, Kenya, 2013-2014', *PLoS Neglected Tropical Diseases*, 10: e0004981.

Maljkovic Berry, I. et al. (2019) 'Global Outbreaks and Origins of a Chikungunya Virus Variant Carrying Mutations Which May Increase Fitness for *Aedes aegypti*: Revelations from the 2016 Mandera, Kenya Outbreak', *The American Journal of Tropical Medicine and Hygiene*, 100: 1249–57.

Martin, D. et al. (2015) 'RDP4: Detection and Analysis of Recombination Patterns in Virus Genomes', *Virus Evolution*, 1: vev003.

Martin, E. et al. (2016) 'Insights into the Molecular Evolution of Dengue Virus Type 4 in Puerto Rico over Two Decades of Emergence', *Virus Research*, 213: 23–31.

McCarthy, D. D., and Wilson, D. B. (1948) 'Dengue in the East African Command; Incidence in Relation to Aedes Prevalence and Some Clinical Features', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 42: 83–8.

Mease, L. E. et al. (2011) 'Seroprevalence and Distribution of Arboviral Infections among Rural Kenyan Adults: A Cross-Sectional Study', *Virology Journal*, 8: 371.

Murray, N. E. A., Quam, M. B., and Wilder-Smith, A. (2013) 'Epidemiology of Dengue: Past, Present and Future Prospects', *Clinical Epidemiology*, 5: 299–309.

Njeru, I. (2013). Dengue/DHF update (33): Asia, Africa, Pacific, ProMED-mail <https://promedmail.org/promed-post/?id=1676860> accessed 6 Mar 2020.

Nurk, S. et al. (2013) 'Assembling Single-Cell Genomes and Mini-Metagenomes from Chimeric MDA Products', *Journal of Computational Biology*, 20: 714–737.

Obonyo, M., Fidhow, A., and Ofula, V. (2018) 'Investigation of Laboratory Confirmed Dengue Outbreak in North-Eastern Kenya, 2011', *PLoS One*, 13: e0198556.

Pascalis, H. et al. (2019) 'The Epidemic of Dengue Virus Type-2 Cosmopolitan Genotype on Reunion Island Relates to Its Active Circulation in the Southwestern Indian Ocean Neighboring Islands', *Heliyon*, 5: e01455.

ProMED-mail (2011). Dengue/DHF update 2011 (39).20111004. 2985, ProMED-mail.

Rambaut, A. et al. (2016) 'Exploring the Temporal Structure of Heterochronous Sequences Using TempEst (Formerly Path-O-Gen)', *Virus Evolution*, 2: vew007.

—— et al. (2018) 'Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7', *Systematic Biology*, 67: 901–4.

ReliefWeb (2017) 'Burkina Faso: Dengue Outbreak—Sep 2017' <https://reliefweb.int/disaster/ep-2017-000162-bfa> accessed 6 Mar 2020.

Ronquist, F. et al. (2012) 'MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space', *Systematic Biology*, 61: 539–42.

Sang, R. C., and Dunster, L. M. (2001) 'The Growing Threat of Arbovirus Transmission and Outbreaks in Kenya: A Review', *East African Medical Journal*, 78: 655–61.

Schmieder, R., and Edwards, R. (2011) 'Quality Control and Preprocessing of Metagenomic Datasets', *Bioinformatics*, 27: 863–4.

Stamatakis, A. (2014) 'RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies', *Bioinformatics*, 30: 1312–3.

Suchard, M. A. et al. (2018) 'Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10', *Virus Evolution*, 4: vey016.

The New Humanitarian (2013). *New Cases of Dengue Fever, Kala-Azar Reported in Kenya's North* <http://www.thenewhumanitarian.org/news/2013/02/19/new-cases-dengue-fever-kala-azar-reported-kenya-s-north> accessed 20 Jan 2020.

Thomas, S. J., and Rothman, A. L. (2015) 'Trials and Tribulations on the Path to Developing a Dengue Vaccine', *Vaccine*, 33: D24–31.

Twiddy, S. S. et al. (2002) 'Phylogenetic Relationships and Differential Selection Pressures among Genotypes of Dengue-2 Virus', *Virology*, 298: 63–72.

Volk, S. M. et al. (2010) 'Genome-Scale Phylogenetic Analyses of Chikungunya Virus Reveal Independent Emergences of Recent Epidemics and Various Evolutionary Rates', *Journal of Virology*, 84: 6497–504.

Waman, V. P. et al. (2016) 'Analysis of Genotype Diversity and Evolution of Dengue Virus Serotype 2 Using Complete Genomes', *PeerJ*, 4: e2326.

Waterhouse, A. M. et al. (2009) 'Jalview Version 2–a Multiple Sequence Alignment Editor and Analysis Workbench', *Bioinformatics*, 25: 1189–91.

Weaver, S. C., and Vasilakis, N. (2009) 'Molecular Evolution of Dengue Viruses: Contributions of Phylogenetics to Understanding the History and Epidemiology of the

Preeminent Arboviral Disease', *Infection, Genetics and Evolution*, 9: 523–40.

Webster, D. P., Farrar, J., and Rowland-Jones, S. (2009) 'Progress towards a Dengue Vaccine', *The Lancet Infectious Diseases*, 9: 678–87.

Wei, K., and Li, Y. (2017) 'Global Evolutionary History and Spatio-Temporal Dynamics of Dengue Virus Type 2', *Scientific Reports*, 7: 45505.

WHO (1993). *Monograph on Dengue/Dengue Haemorrhagic Fever*. Thongcharoen P. New Delhi, India: WHO Regional Office for South-East Asia.

—— (1999). *Prevention and Control of Dengue and Dengue Haemorrhagic Fever: Comprehensive Guidelines*, World Health Organization: No. 29, pp. 21–134. New Delhi, India: WHO Regional Office for South-East Asia.

WHO (2011). *Horn of Africa Crisis: November 2011 Update*, World Health Organization <https://www.who.int/hac/crises/horn_of_africa/who_hoa_rep_nov2011.pdf?ua=1> accessed 6 Mar 2020.

WHO (2014). *Dengue and Severe Dengue (No. WHO-EM/MAC/032/E)*, World Health Organization. Cairo, Egypt: WHO Regional Office for the Eastern Mediterranean.

—— (2016). *Emergencies Preparedness, Response: Dengue Fever* – Burkina Faso, World Health Organization <https://www.who.int/csr/don/18-november-2016-dengue-burkina-faso/en/> accessed 6 Mar 2020.

—— (2017). *Weekly Bulletin on Outbreaks and Other Emergencies: Week 28: 08-14 July 2017*, World Health Organization <https://apps.who.int/iris/bitstream/handle/10665/255895/OEW28-81472017.pdf?> accessed 6 Mar 2020.