

Spatiotemporal patterns of Rift Valley fever virus in Africa: a retrospective genomic epidemiology and phylodynamic modelling study



John Juma, Houriyah Tegally, Samson L Konongoi, Kristina Roesel, Rosemary Sang, Tulio de Oliveira, M Kariuki Njenga, Alan Christoffels, Bernard Bett, Simon Dellicour, Samuel O Oyola



Summary

Background Rift Valley fever virus (RVFV) is a mosquito-borne zoonotic pathogen causing outbreaks in humans and ruminants across Africa and the Arabian Peninsula. Originally restricted to the Great Rift Valley, RVFV has expanded geographically, prompting its classification by WHO as a pathogen of pandemic potential. We investigated the evolutionary and spatial dynamics of RVFV across Africa.

Methods We used genomic data generated at the International Livestock Research Institute Nairobi genomic laboratory (BioProject PRJNA1106221) and combined with publicly available datasets retrieved from the National Center for Biotechnology (NCBI) GenBank nucleotide database. In retrieving RVFV genome sequences from the NCBI GenBank, we applied the search terms “Rift Valley fever virus segment L AND 6404[SLEN]”, “Rift Valley fever virus segment M AND 3885[SLEN]”, and “Rift Valley fever virus segment S AND 1520:1690[SLEN]” for L (Large), M (Medium), and S (Small) segments, respectively. For sequences without additional spatiotemporal information, we searched PubMed to extract the associated sequence metadata. We performed molecular clock analysis, phylogenetic inference, phylodynamic modelling (continuous phylogeographic reconstruction), and landscape phylogeography on the three RVFV genome segments (L, M, and S). We aimed to assess evolutionary rates, dispersal patterns, and environmental drivers. Focus was placed on lineage C, the most widely distributed variant.

Findings The global dataset used in this study consisted of large (n=236), medium (n=237), and small (n=247), which were further filtered to exclude potential reassortants and vaccine strains. Genome sequences retrieved from NCBI GenBank database comprised large (n=180), medium (n=184), and small (n=202). The genome sequences from retrospective human and livestock isolates comprised large (n=56), medium (n=53), and small (n=45) collected in Burundi (2018), Kenya (2007, 2018, 2019, 2021, and 2022), and Rwanda (2018 and 2022). Our dataset revealed that RVFV exhibited low overall genetic diversity. Lineage C, however, showed evidence of active evolution, with substitution rates ranging from 3.58×10^{-4} to 9.76×10^{-4} substitutions per site per year. This lineage probably originated in Zimbabwe in the mid-1970s and has since expanded across eastern and southern Africa. Phylogeographic reconstructions revealed rapid spread, with diffusion coefficients exceeding 50 000 km² per year.

Interpretation Lineage C appears capable of establishing endemic transmission in new regions, with ongoing diversification observed during interepidemic periods. These observations reinforce the value of continuous genomic surveillance, particularly during cryptic transmission phases when adaptive mutations might emerge. Although further evidence is needed, observed trends in climate variability and land-use change point to the potential benefit of targeted surveillance in settings that could be at increased risk, including urban centres and wetlands.

Funding This work was supported by the German Federal Ministry for Economic Cooperation and Development, the Rockefeller Foundation, and the Africa Centres for Disease Control and Prevention.

Copyright © 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

Introduction

Rift Valley fever (RVF) is a mosquito-borne zoonotic disease prioritised by WHO and the African Union for research and development due to its pandemic potential and significant public health impact.¹ The disease is characterised by high mortality and abortion rates in domestic ruminants, along with substantial morbidity in humans. The causative agent, Rift Valley fever virus (RVFV), is a segmented RNA virus

(L, M, and S segments) of the *Phlebovirus* genus in the Phenuiviridae family.² Genome reassortment among segments has been previously reported.³

RVFV is recognised as both a zoonotic and biosecurity threat.⁴ Although there is no licensed human vaccine, a promising live-attenuated candidate (hRVFV-4s) is in phase 1 clinical trials in Belgium.⁵ Available veterinary vaccines in Africa remain limited due to safety concerns,

Lancet Microbe 2026;
7: 101251

Published Online February 18,
2026

<https://doi.org/10.1016/j.lanmic.2025.101251>

International Livestock Research Institute, Nairobi, Kenya (J Juma PhD, S L Konongoi BVetMed, K Roesel PhD, B Bett PhD, S O Oyola PhD); South African National Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa (J Juma, Prof A Christoffels PhD, S O Oyola); Centre for Epidemic Response and Innovation, School of Data Science and Computational Thinking, Stellenbosch University, Stellenbosch, South Africa (H Tegally PhD, Prof T de Oliveira PhD); Centre for Virus Research, Kenya Medical Research Institute, Nairobi, Kenya (S L Konongoi, Prof R Sang PhD); Washington State Global Health Program—Kenya, Washington State University, Nairobi, Kenya (Prof M K Njenga PhD); Spatial Epidemiology Laboratory, Université Libre de Bruxelles, Brussels, Belgium (S Dellicour PhD); Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory for Clinical and Epidemiological Virology, KU Leuven—University of Leuven, Leuven, Belgium (S Dellicour)

Correspondence to:
Dr Samuel O Oyola,
International Livestock Research Institute, Nairobi 00100, Kenya
S.Oyola@cgiar.org

Research in context

Evidence before this study

We searched PubMed from Jan 1, 1930, to Dec 20, 2023, using the terms “Rift Valley fever virus” and “phylogeography”. Of six relevant articles, three focused on Rift Valley fever virus (RVFV) phylogeography in Africa, the Arabian Peninsula, and Indian Ocean islands. These studies used Bayesian phylogeographic methods, such as discrete trait analysis, to infer transmission routes. However, none investigated the integration of fine-scale environmental data with genomic analyses, particularly focusing on lineage C—the predominant RVFV lineage in Africa.

Added value of this study

Using a comprehensive dataset, we performed spatiotemporal phylodynamic analysis of RVFV across Africa. We identified that

lineage C is evolving at a faster rate than other lineages and is diversifying into sublineages. Our analysis traced its origin to Zimbabwe in the mid-1970s, followed by regional dissemination, including into Kenya.

Implications of all the available evidence

Our findings offer new insights into RVFV lineage C circulation across Africa, including Indian Ocean islands (Mayotte and Madagascar) and Arabian Peninsula settings. Although earlier studies outlined regional virus spread, they lacked integration of environmental correlates. Our work demonstrates the utility of combining genomic and environmental data in a phylogeographic framework to better understand RVFV dispersal dynamics and inform surveillance strategies.

particularly in pregnant animals, where they have been associated with teratogenic outcomes.⁶

Since the first detection in Kenya in 1931, at least 13 countries across Africa and beyond, including Saudi Arabia and Yemen, have reported outbreaks.⁷ Recently, RVF activity has also been reported in Burundi and Rwanda.^{8,9} These outbreaks often coincide with above-average rainfall, frequently linked to El Niño Southern Oscillation events.¹⁰ Transmission during interepidemic periods (IEPs) is thought to occur via vertical transmission in mosquito eggs deposited in shallow depressions (ie, dambos), which hatch under favourable climatic conditions.¹¹ However, the drivers of outbreak onset and viral persistence during IEPs remain incompletely understood. This study aimed to investigate the evolutionary characteristics, transmission dynamics, and potential environmental and climatic factors associated with the spread of RVFV in Africa.

Methods

Study design and samples

In this retrospective genomic epidemiology study, we used RVFV genome sequences from humans, livestock, and mosquito vectors. We applied targeted viral enrichment and whole-genome sequencing to generate genomic data from archived samples collected during sporadic RVF outbreaks. We analysed RVFV genomic data generated at the International Livestock Research Institute Nairobi genomic laboratory (BioProject PRJNA1106221) and supplemented it with publicly available sequences from the National Center for Biotechnology Information (NCBI) GenBank nucleotide database (appendix pp 1–3). In retrieving RVFV genome sequences from the NCBI GenBank, we applied the search terms “Rift Valley fever virus segment L AND 6404[SLEN]”, “Rift Valley fever virus segment M AND 3885[SLEN]”, and “Rift Valley fever virus segment S AND 1520:1690[SLEN]” for L, M, and S segments, respectively. We included sequences with valid sampling or collecting dates (dating from 1944 to 2022). Because reassortment might result in divergent

evolutionary histories for the L, M, and S genome segments,³ analyses were performed separately for each segment.³ Only sequences with greater than 90% genome coverage were retained.

Procedures

Multiple sequence alignments were performed using MAFFT (version 7.475),¹² followed by manual inspection and primer site trimming using AliView (version 1.28).¹³ Best-fitting nucleotide substitution models were identified with jModelTest (version 0.1.7)¹⁴ and used in downstream phylogenetic analyses. Maximum likelihood phylogenies were reconstructed using IQ-TREE (version 2.0.3)¹⁵ with 1000 ultrafast bootstrap replicates.

We assessed temporal signal by regressing root-to-tip distances against sampling dates using TempEst (version 1.5.3). Datasets showing sufficient temporal signal were analysed using BEAST (version 1.10.4).¹⁶ Continuous phylogeographic reconstructions were performed under a relaxed random walk diffusion model with branch-specific heterogeneity in diffusion velocity modelled using a Cauchy distribution. A GTR+I+G4 substitution model, an uncorrelated log-normal relaxed molecular clock, and a skygrid coalescent prior were implemented. Markov chain Monte Carlo analyses ran for 1.5×10^9 steps (L), 1.2×10^9 steps (M), and 6.0×10^8 steps (S), sampling every 150 000, 120 000, and 60 000 steps, respectively. BEAGLE (version 3.1.2)¹⁷ was used to accelerate likelihood computation. Convergence and effective sample sizes (>200) were evaluated with Tracer (version 1.7.2).¹⁸ Maximum clade credibility trees were summarised using TreeAnnotator.¹⁶

To visualise and quantify viral dispersal, 300 post-burn-in trees were analysed with the R package *seraphim*,¹⁹ from which we estimated weighted diffusion coefficients²⁰ and wavefront distances. We assessed the impact of environmental factors on lineage diffusion velocity in a four-step procedure and tested associations between viral lineage dispersal locations and ecological variables (appendix pp 19–21).

See Online for appendix

Role of the funding source

The funders of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the manuscript.

Results

The global RVFV dataset used in the study comprised 236, 237, and 247 genome sequences for the L, M, and S segments, respectively. These genomes were further categorised per segment as large (NCBI n=180, outbreak n=56), medium (NCBI n=184, outbreak n=53), and small (NCBI n=202, outbreak n=45; table). Available sequences primarily originated from Africa and Saudi Arabia, with the largest contributions from South Africa (n=93) and Kenya (n=49; figure 1A).

Genotype assignment²¹ indicated that lineage C is the most prevalent (n=115; 48.52%) RVFV lineage (figure 1B; appendix pp 5–17). This lineage is predominantly found in eastern Africa, whereas lineage H is more common in southern Africa, notably during the 2010–11 outbreak in South Africa. Lineage A has been reported in northern Africa, particularly during the 1977 outbreak in Egypt. The ancestral lineage K was only detected in Uganda and Zimbabwe, with recent sequences from Uganda (2018) suggesting continued endemicity (figure 1B).

Root-to-tip regression analysis showed a strong temporal signal between genetic divergence and sampling dates (figure 2A; appendix pp 21–22), supporting the suitability of these data for molecular clock-based evolutionary analyses.

Molecular clock analysis estimated the time to the most recent common ancestor (tMRCA) of RVFV to be around 1918 (95% highest posterior density [HPD] 1902–30). The estimated tMRCAs for lineages A, C, and H were 1971 (1968–73), 1970 (1966–75), and 2007 (2006–08), respectively (figure 2B). The evolutionary rate of the complete medium segment was 4.56×10^{-4} substitutions per site per year (95% HPD 3.88×10^{-4} – 5.32×10^{-4}), consistent with rates observed in other RNA viruses.

Skygrid reconstruction revealed a steady increase in the effective viral population size between the early 1900s and 1930s, followed by episodic expansions approximately every 10–15 years (figure 2C). Phylogenetic analyses showed consistent topologies across all three genome segments, with most sequences clustering into lineages C and H (figure 2D, appendix p 21).

To account for rate variation, separate molecular clock analyses were done for each segment. For lineage C, tMRCA estimates were 1965 (95% HPD 1952–74) for the large segment, 1965 (1949–75) for the medium segment, and 1971 (1962–75) for the small non-structural segment (figure 3A; appendix pp 23–25). Our analyses suggest that lineage C originated in Zimbabwe in 1976 and was later introduced independently into Kenya (1983), Madagascar (1990s), and the Arabian Peninsula and South Africa (2000s; figure 3A, B; appendix pp 23–25).

Phylogenetic clustering based on branching time and tMRCA estimates revealed distinct sublineages within lineage C, including C.1, C.2.1, C.1.1, C.1.2, and C.2.2, many

	Large-segment genome sequences (n=236)	Medium-segment genome sequences (n=237)	Small-segment genome sequences (n=247)
Year of collection	2010 (1944–2022)	2010 (1944–2022)	2010 (1944–2022)
Host			
Humans	128 (54%)	129 (54%)	126 (51%)
Bat	1 (<1%)	2 (1%)	1 (<1%)
Buffalo	5 (2%)	4 (2%)	6 (2%)
Camel	1 (<1%)
Cow	74 (31%)	73 (31%)	75 (30%)
Goat	2 (1%)	2 (1%)	..
Mosquito	17 (7%)	20 (8%)	21 (9%)
Sheep	9 (4%)	7 (3%)	16 (6%)
Springbok	1 (<1%)
Lineage			
A	9 (4%)	10 (4%)	11 (4%)
B	1 (<1%)	1 (<1%)	1 (<1%)
C	114 (48%)	115 (49%)	141 (57%)
D	1 (<1%)	1 (<1%)	1 (<1%)
E	5 (2%)	2 (1%)	2 (1%)
F	1 (<1%)	1 (<1%)	1 (<1%)
G	2 (1%)	6 (3%)	5 (2%)
H	82 (35%)	81 (34%)	59 (24%)
I	2 (1%)	2 (1%)	2 (1%)
J	1 (<1%)	1 (<1%)	1 (<1%)
K	6 (3%)	3 (1%)	6 (2%)
L	5 (2%)	6 (3%)	7 (3%)
M	..	1 (<1%)	3 (1%)
N	6 (3%)	6 (3%)	5 (2%)
O	1 (<1%)	1 (<1%)	2 (1%)
Source			
NCBI	180 (76%)	184 (78%)	202 (82)
ILRI collection	56 (24%)	53 (22%)	45 (18%)
Data are median (IQR) or n (%). NCBI=National Center for Biotechnology Information. ILRI=International Livestock Research Institute.			
Table: Distribution of Rift Valley fever virus genomic sequences, viral lineages, and data sources included in the analysis			

of which emerged during or after the 1997 and 2006–07 outbreaks in Kenya (figure 2A, B). Lineage C exhibited evolutionary rates of 3.58×10^{-4} (95% HPD 2.59×10^{-4} – 4.55×10^{-4}), 5.07×10^{-4} (3.35×10^{-4} – 6.57×10^{-4}), and 9.76×10^{-4} (6.91×10^{-4} – 1.28×10^{-3}) substitutions per site per year for the large, medium, and small non-structural segments, respectively (figure 3C).

Phylogeographic reconstruction of RVFV lineage C revealed its early spread across several locations in Zimbabwe, including Kadoma (formerly Gatooma), Harare (formerly Salisbury), and Chinhoyi (formerly Sinoia), with the most probable origin dating to 1965 (95% HPD 1949–75; figure 4A; appendix pp 26–27). The virus subsequently dispersed to Ruiru, a town in central Kenya, in around 1983 (1955–77). A separate introduction from Zimbabwe into Madagascar was inferred in

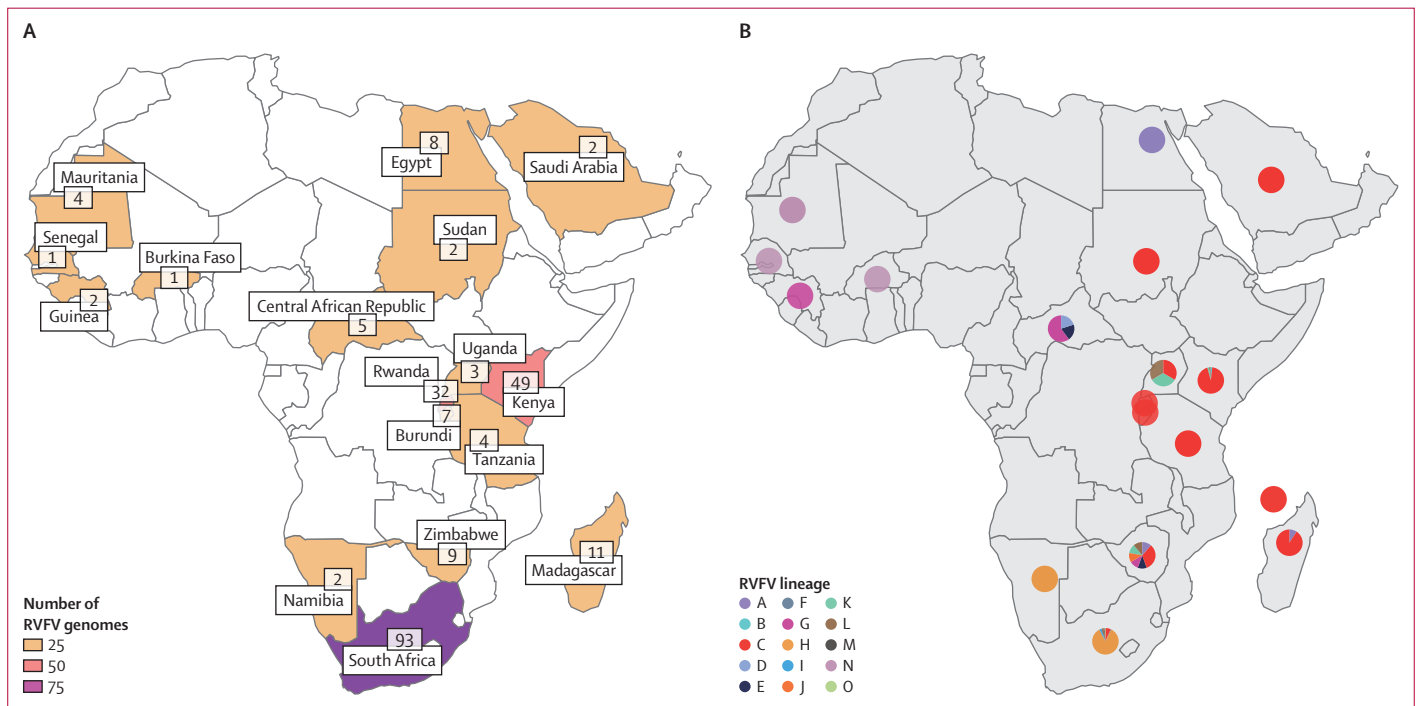


Figure 1: Distribution of RVFV genomic sequences

Genomic sequence data for the medium segment retrieved from NCBI GenBank database and classified at the lineage level. (A) The distribution of genomic sequences consisting of public (NCBI as of Dec 20, 2023) and locally generated consensus genomes of RVFV. (B) Scatterpie plot depicting occurrence of RVFV lineages per country. The radii of the circles correspond to the number of isolates (log-transformed) identified in each country. RVFV=Rift Valley fever virus. NCBI=National Center for Biotechnology Information.

1991 (1986–91), possibly coinciding with an unobserved introduction into South Africa during this period. The earliest evidence of lineage C circulation in South Africa dates to approximately 1999.

Another notable dissemination event was the introduction of lineage C into the Arabian Peninsula in the early 2000s (95% HPD 1998–2001), following the major eastern African outbreak in 1997. From the mid-2000s, lineage C persisted in Kenya and spread into neighbouring eastern African countries, including Tanzania (2007, 2004–07) and Sudan (2010, 2008–10). Additional evidence supports a transmission event from Madagascar to Mayotte in 2007 (2005–07).

Lineage C also appeared in Uganda in 2015 (95% HPD 2010–15), potentially introduced via Sudan, where outbreaks were documented in 2007. Recent outbreaks in Rwanda (Rulindo and Ngoma) in 2018 (2017–18) were likely seeded by viral migration from Kenya via Uganda. Our data further suggest that Rwanda might have subsequently served as a source for viral export to Uganda (2017, 2015–20) and Burundi (2022, 2020–22).

Taken together, our analysis identifies Zimbabwe as the likely origin of lineage C, with Kenya playing a major role in subsequent eastern African dissemination. South African outbreaks appear to result from distinct introduction events originating independently from both Zimbabwe and Kenya (figure 4A; appendix pp 26–27).

To quantify the spatial dynamics of viral spread, dispersal statistics were estimated, including the weighted diffusion coefficient (WDC) and the evolution of maximal wavefront distance. WDC, being less sensitive to sampling bias than velocity-based metrics, was used as the primary measure of lineage movement.²⁰ Lineage C has spread over a geographical range exceeding 6000 km (95% HPD 0–7621) from its epidemic origin (figure 4B; appendix pp 26–27). The highest diffusion rate was observed in the small non-structural segment (83 049 km²/year, 95% HPD 58 697–109 040), compared with the medium (53 988 km²/year, 34 621–74 221) and large (56 315 km²/year, 39 972–72 781) segments.

The wavefront analysis revealed a stepwise expansion pattern, with gradual increases in dispersal distance followed by plateau phases. Notable surges in wavefront distance occurred around 1975, 1991, 1997, and 2007—timepoints that correspond closely with known outbreak events in Zimbabwe, Madagascar, and eastern Africa (1997–98 and 2006–07).

Estimates of the effective viral population size showed consistent demographic trends across all three genomic segments (figure 4C; appendix pp 26–27). After a decline in the 1980s, the effective population size increased during the 1990s, peaking in 1997–98. A decline followed in the early 2000s, with another peak observed during the 2006–07 outbreaks. Since then, a gradual decrease in effective population size has been evident up to 2018.

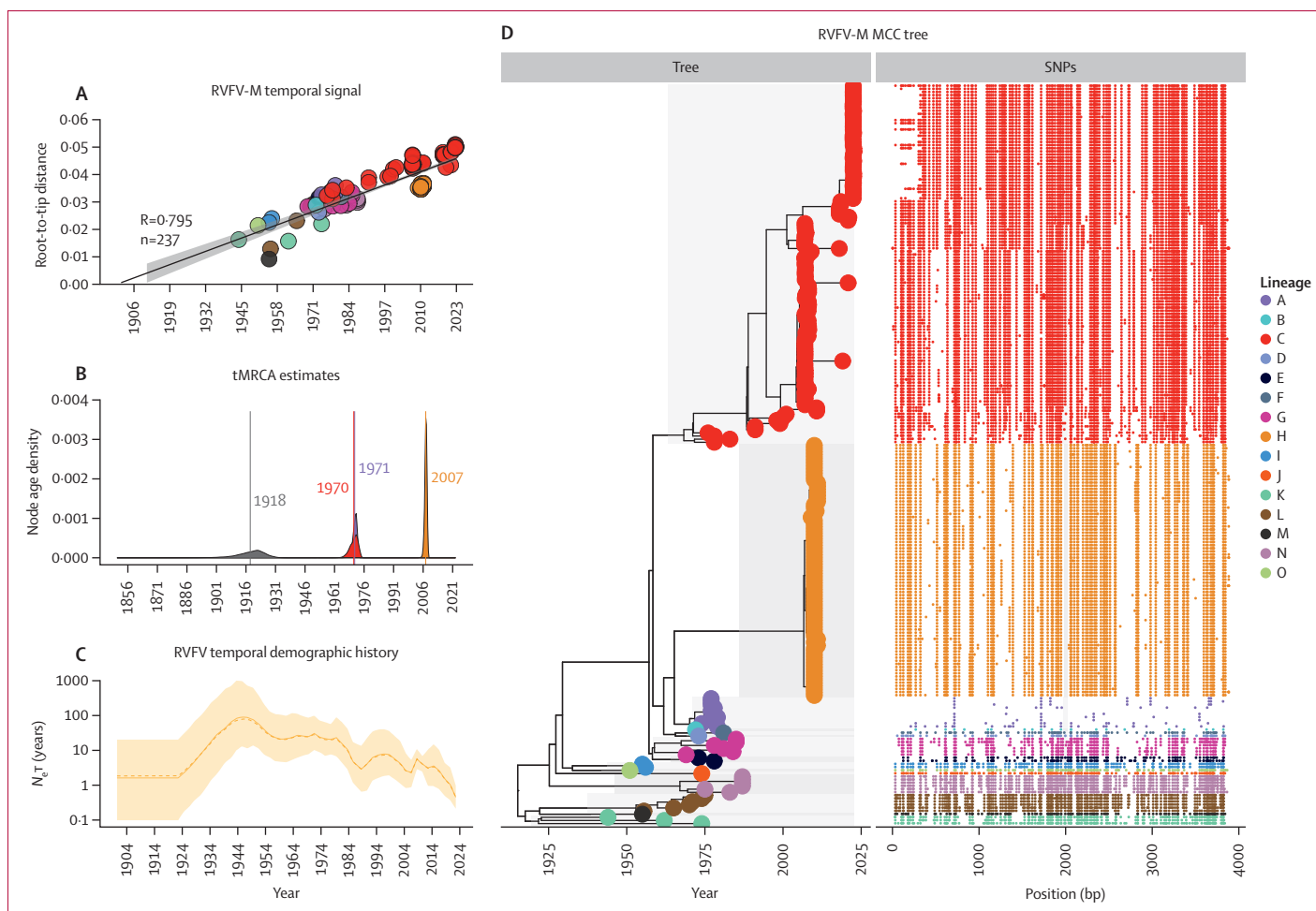


Figure 2: Phylogenetic analysis of RVFV genomic sequences

Genomic sequence data assigned to RVFV lineages and time-scaled phylogenetic tree inference. Maximum likelihood phylogenetic trees were constructed from the complete medium segment glycoprotein sequences ($n=237$) using GTR + I + G4 as the substitution model with 1000 replicates as bootstrapping procedure. (A) The temporal signal of the medium-segment genetic sequences assessed by regression of the root-to-tip genetic distance and time (years). (B) Estimation of the time of the most recent ancestors for lineages A, C, and H. (C) Population dynamics of RVFV under a skygrid coalescent model representing the estimated viral effective population sizes over the past 100 years. (D) MCC tree retrieved from the Bayesian phylogenetic inference based on all the genetic sequences, here highlighting the clustering of the different lineages of the virus and associated SNP for each lineage on the right panel. Each tip on the tree represents a virus sample and is coloured according to the assigned lineage. RVFV=Rift Valley fever virus. M=medium. SNP=single nucleotide polymorphism. MCC=maximum clade credibility. tMRCA=time to the most recent common ancestor.

We investigated whether the locations of RVFV lineage dispersal events, defined as phylogenetically inferred movements between geographical locations, were associated with specific environmental conditions. When considering only results consistently supported across all three genomic segments, sampled viral lineages were more likely to circulate in cropland and urban landscapes characterised by relatively higher human population densities (appendix p 28). However, as more than half of the tree node positions analysed for environmental associations corresponded to actual sampling locations, these findings are inherently sensitive to sampling effort and spatial biases. As such, they should be interpreted as descriptive of the environmental conditions associated with sampled viral lineages, rather than definitive drivers of viral movement.

We also examined the association between environmental factors and the diffusion velocity of RVFV lineages using

continuous phylogeographic reconstructions. Only one strong and consistent association was observed: primary non-forest areas appeared to act as a resistance factor to viral dispersal in the analysis of the small (non-structural) genomic segment. Notably, this association was not observed in the analyses of the large and medium genomic segments, which did not support primary non-forest areas as impediments to viral diffusion. These inconsistencies across genomic segments suggest limited or context-dependent influence of environmental barriers on RVFV lineage spread (appendix pp 30–33).

Discussion

In this study, we applied genomic sequencing and phylogenetic approaches to investigate the evolutionary and dispersal dynamics of RVFV across Africa. Although earlier studies assessed the geographical spread of RVFV,^{22,23} a

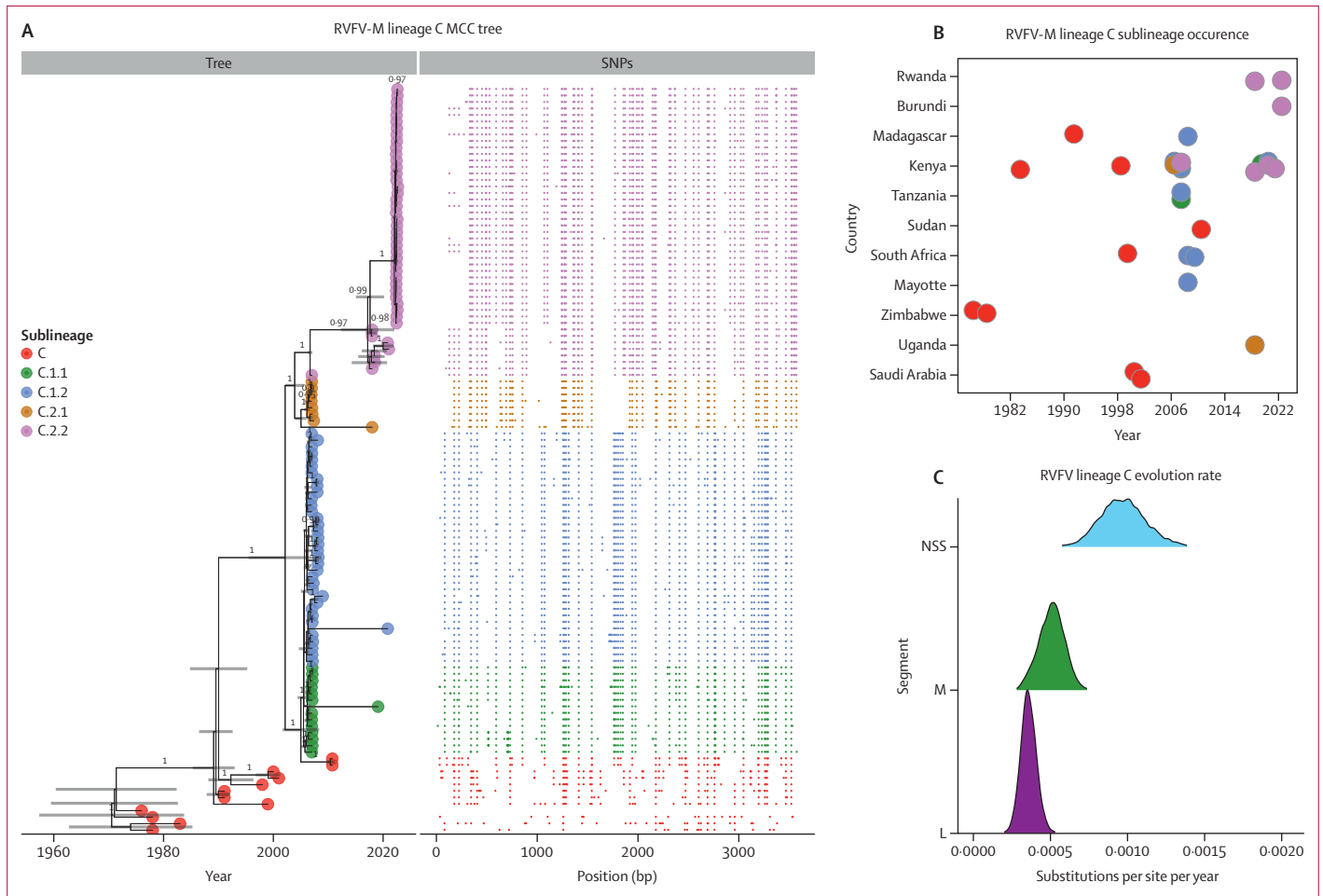


Figure 3: Evolutionary analysis of RVFV lineage C

(A) MCC tree of RVFV lineage C based on glycoprotein genomic sequences inferred using Bayesian phylogenetics. Posterior probabilities are shown on branches, with 95% HPD intervals represented as grey bars. Tree annotations include host metadata and sublineage-specific SNP profiles. (B) Spatiotemporal distribution of lineage C sublineages across Africa and the Arabian Peninsula from the 1970s (Zimbabwe) to 2022 (Rwanda and Burundi). (C) Posterior distribution of the estimated evolutionary substitution rates for lineage C. RVFV=Rift Valley fever virus. MCC=maximum clade credibility. HPD=highest posterior density. SNP=single nucleotide polymorphism.

comprehensive analysis integrating georeferenced genomic data with environmental factors was lacking. Our findings highlight the continued circulation and geographical expansion of lineage C, a key driver of cryptic transmission in the region.

Although clinical surveillance reports are abundant, RVFV genomic data remain sparse. By generating new sequence data and contextualising them within the broader genomic landscape, we revealed extensive viral circulation across Africa. Our phylogenetic analyses confirmed lineage-specific geographical clustering—lineage A predominates in northern Africa, lineage C in eastern Africa, and lineage H in southern Africa, with occasional detection in west Africa. Notably, the presence of lineages in previously RVF-free areas suggests that the virus can establish itself in new territories under favourable conditions.

Molecular clock analyses estimated the tMRCA of RVFV around 1918 (95% HPD 1902–30), consistent with earlier

studies. This period coincides with European colonial expansion and livestock introduction to African highlands, potentially contributing to RVFV emergence.²⁴ The virus exhibits a low evolutionary rate (4.56×10^{-4} substitutions per site per year), probably due to constraints imposed by its dual-host (vector and mammal) lifecycle—supporting the double filter hypothesis.²⁵ Nevertheless, the detection of distinct sublineages within lineage C suggests that adaptive evolutionary processes might still be at play.

Our analysis showed that genes involved in replication (RdRp), envelope formation (Gn/Gc), and immune evasion (NSS) have elevated substitution rates, potentially driving phenotypic diversification. Sublineages C.1.1 and C.1.2 (linked to the 1997–98 Kenya outbreak), and C.2.1 and C.2.2 (linked to the 2006–07 outbreak), point to ongoing evolution within lineage C. These patterns hypothesise that mutation accumulation during IEPs might contribute to cryptic transmission. Targeted genomic and

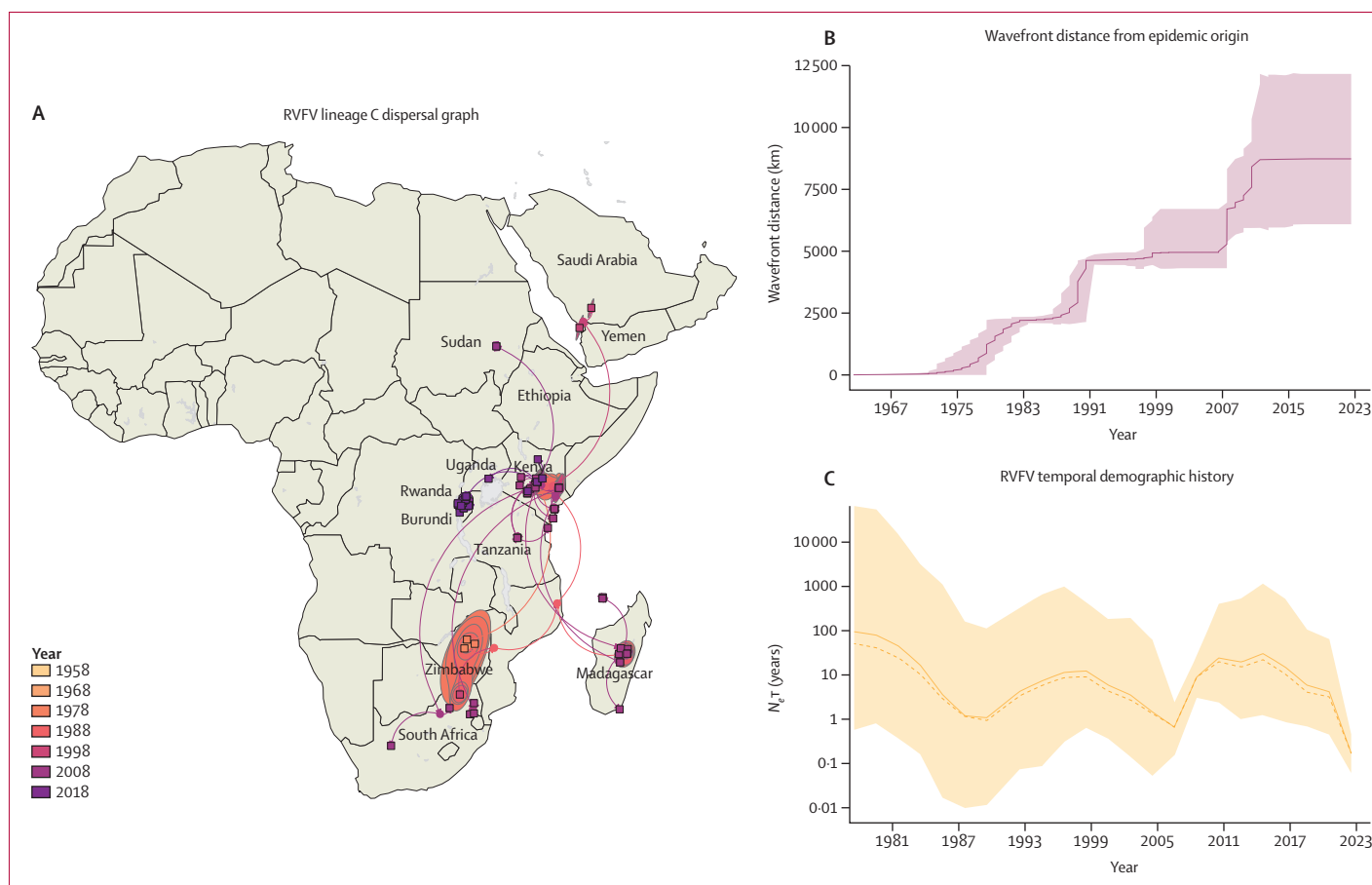


Figure 4: Spatiotemporal dispersal dynamics of RVFV lineage C in Africa inferred using medium segment sequences

(A) Dispersal history of RVFV lineage C summarised by a MCC tree retrieved and annotated from 1000 posterior trees sampled from the posterior distribution of a continuous phylogeographic analysis. Nodes of the trees are coloured according to the time of occurrence, and oldest nodes are plotted on top of youngest ones. Internal nodes are shown by circles, while tip nodes are square. (B) RVFV lineage C dispersal statistics showing the evolution of the maximal wavefront distance from the epidemic origin (median 447 km, 95% HPD 0–7621). (C) The effective virus population size (log-transformed) estimated over the past 50 years (median 7, 95% HPD 1–91). The shaded areas correspond to the 95% HPD intervals in all panels, the solid lines represent the median values of estimates of interest (maximal wavefront distance and effective viral population size), and the dashed line in the effective population size indicates the median estimate. RVFV=Rift Valley fever virus. MCC=maximum clade credibility. HPD=highest posterior density.

serological surveillance in livestock and wildlife is essential for early detection during these silent transmission cycles.

Our continuous phylogeographic analyses suggest that lineage C likely emerged in Zimbabwe (1949–75), followed by spread to Kenya, and subsequently to neighbouring countries. The virus has since expanded into the Arabian Peninsula, Mayotte, and Madagascar, underscoring its capacity to cross ecological and geopolitical barriers.^{22,23} Dispersal patterns are characterised by 10–15-year intervals, coinciding with outbreak events and marked increases in wavefront distance, effective population size, and diffusion coefficients.

Demographic modelling indicated large virus population sizes during major outbreaks, which probably facilitated adaptation through natural selection.²⁵ Population bottlenecks during periods of reduced transmission, possibly influenced by vaccination campaigns in the 1980s and 2016–17, might have temporarily limited viral evolution.²⁶

Lineage C shows high spatial diffusivity (>50 000 km²/year), similar to other vector-borne viruses, such as West Nile virus and lumpy skin disease virus.²⁰ International livestock trade, transboundary animal movement, and nomadic pastoralism probably play crucial roles in the virus's spread across the Horn of Africa.²⁷

Landscape phylogeographic analyses suggested preferential viral circulation in croplands and urban areas with higher population density, although these findings might reflect sampling bias. Primary non-forested areas might act as barriers to viral diffusion, but this effect was only consistently observed for one genomic segment.

There are limitations to be considered in interpreting the current results. This study is limited by the paucity of RVFV genomic data, particularly from regions outside Africa and the Arabian Peninsula, which hinders a full understanding of its global spread. Sampling efforts were concentrated in areas with higher population densities, introducing potential bias in the analysis of viral distribution. The

cryptic transmission during IEPs could lead to undetected genetic diversity. Additionally, gaps in genomic data, especially from the 1970s to 1980s, might affect the accuracy of conclusions about early virus spread. The study also relies on model assumptions in molecular clock analyses, which might not fully account for all evolutionary variables. These limitations highlight the need for broader geographical coverage, improved surveillance in under-sampled areas, and more comprehensive environmental data to refine our understanding of RVFV evolution and dispersal.

Overall, our study provides critical genomic insights into the evolution and spread of RVFV, with a focus on lineage C. This study shows that this lineage is undergoing ongoing evolution, expanding geographically, and diversifying into new sublineages. These findings underscore the need for routine genomic surveillance, especially during IEPs when selection pressures might shape viral fitness. The ability of RVFV to emerge in new regions, facilitated by livestock movement and changing environments, presents an ongoing challenge to public and animal health systems. Integrating genomic, spatial, and environmental data can inform targeted surveillance and more effective control strategies for RVF. A key contribution of this study is the demonstration that genomic epidemiology is a powerful surveillance tool, capable of revealing transmission dynamics and evolutionary patterns that might remain undetected using classical epidemiological approaches alone.

Contributors

JJ and SOO designed the study. JJ and SLK did the experiments, including molecular screening, and produced RVFV genomic data. JJ, SD, and HT verified the metadata, curated the metadata, and analysed the data. HT, SD, RS, KR, BB, TdO, AC, MKN, and SOO helped in the interpretation of the data. JJ wrote the initial draft of the manuscript. SOO reviewed and edited the manuscript and supervised the project. All authors commented on and revised the manuscript. All authors had full access to all the data in the study and accept responsibility for the decision to submit for publication.

Declaration of interests

We declare no competing interests.

Data sharing

Raw sequence data are available in the sequence read archive (SRA) database with a bioproject accession PRJNA1106221. Sequences generated as part of the study have the accessions OR780613-OR780629, PP746415-PP746471, PP747229-PP747243, and PP746304-PP746414. All BEAST output files (logs and trees) are available on zenodo (<https://doi.org/10.5281/zenodo.12703880>) and all custom code and associated files are available on the GitHub repository (<https://github.com/ajodeh-juma/rvf-phylogeog>).

Acknowledgments

This work was supported by the Boosting Uganda's Investments in Livestock Development project, led by the International Livestock Research Institute and funded by the German Federal Ministry for Economic Cooperation and Development (grant number BMZ001201). Additional funding was provided by the Rockefeller Foundation and the Africa Centres for Disease Control and Prevention through a subgrant awarded to SOO. SOO and JJ are also supported by the Global Health EDCTP3 Joint Undertaking and its members, as well as the Bill & Melinda Gates Foundation (grant agreement number 101103171). SD acknowledges support from the Fonds National de la Recherche Scientifique (Belgium;

grant number F.4515.22), the Research Foundation—Flanders (Belgium; grant number G098321N), and the European Union Horizon 2020 projects MOOD (grant number 874850) and LEAPS (grant number 101094685). MKN received support from the US National Institute of Allergy and Infectious Diseases under grant number U01AI151799 through the Centre for Research in Emerging Infectious Diseases-East and Central Africa. TdO and HT were supported by the Rockefeller Foundation (HTH 017), the US National Institutes of Health (U01 AI151698) through the United World Antiviral Research Network, the Global Health EDCTP3 Joint Undertaking and its members, the Bill & Melinda Gates Foundation (grant number 101103171), the Health Emergency Preparedness and Response Umbrella Program managed by the World Bank Group (TF0B8412), the Medical Research Foundation (MRF-RG-ICCH-2022-100069), the Wellcome Trust (grant number 228186/Z/23/Z), and the Novo Nordisk Foundation (grant number NNF24OC0094346).

Editorial note: The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

References

- Mehand MS, Al-Shorbaji F, Millett P, Murgue B. The WHO R&D Blueprint: 2018 review of emerging infectious diseases requiring urgent research and development efforts. *Antiviral Res* 2018; **159**: 63–67.
- Adams MJ, Lefkowitz EJ, King AMQ, et al. Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2017). *Arch Virol* 2017; **162**: 2505–38.
- Sall AA, Zanotto PMA, Sene OK, et al. Genetic reassortment of Rift Valley fever virus in nature. *J Virol* 1999; **73**: 8196–200.
- Rabinowitz P, Gordon Z, Chudnov D, et al. Animals as sentinels of bioterrorism agents. *Emerg Infect Dis* 2006; **12**: 647–52.
- LARISSA Consortium. Novel Rift Valley fever vaccine dosed in phase I human trial. 2022. <https://www.larissa.online/post/novel-rift-valley-fever-vaccine-dosed-in-phase-i-human-trial> (accessed Oct 11, 2022).
- Botros B, Omar A, Elian K, et al. Adverse response of non-indigenous cattle of European breeds to live attenuated Smithburn Rift Valley fever vaccine. *J Med Virol* 2006; **78**: 787–91.
- Nanyingi MO, Munyua P, Kiama SG, et al. A systematic review of Rift Valley Fever epidemiology 1931-2014. *Infect Ecol Epidemiol* 2015; **5**: 28024.
- FluTrackers.com. Rwanda - outbreaks of Rift Valley fever in cattle and possibly human (OIE, April 09, 2022). FluTrackers News and Information. 2022. <https://flutrackers.com/forum/forum/emerging-diseases-other-health-threats-alphabetical-i-thru-z/rift-valley-fever/945819-rwanda-outbreaks-of-rift-valley-fever-in-cattle-and-possibly-human-oie-april-09-2022> (accessed Jan 19, 2023).
- UN Office for the Coordination of Humanitarian Affairs. The government of Burundi, FAO, and WHO continue to combat Rift Valley fever (RVF). 2022. <https://reports.unocha.org/en/country/burundi/card/7q7NfsPS0N/> (accessed Jan 19, 2023).
- Linthicum KJ, Anyamba A, Tucker CJ, Kelley PW, Myers MF, Peters CJ. Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya. *Science* 1999; **285**: 397–400.
- Linthicum KJ, Britch SC, Anyamba A. Rift Valley fever: an emerging mosquito-borne disease. *Annu Rev Entomol* 2016; **61**: 395–415.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013; **30**: 772–80.
- Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 2014; **30**: 3276–78.
- Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 2012; **9**: 772.
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015; **32**: 268–74.
- Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 2018; **4**: vey016.

- 17 Ayres DL, Cummings MP, Baele G, et al. BEAGLE 3: improved performance, scaling, and usability for a high-performance computing library for statistical phylogenetics. *Syst Biol* 2019; **68**: 1052–61.
- 18 Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 2018; **67**: 901–04.
- 19 Dellicour S, Rose R, Faria NR, Lemey P, Pybus OG. SERAPHIM: studying environmental rasters and phylogenetically informed movements. *Bioinformatics* 2016; **32**: 3204–06.
- 20 Dellicour S, Bastide P, Rocu P, et al. How fast are viruses spreading in the wild? *PLoS Biol* 2024; **22**: e3002914.
- 21 Juma J, Fonseca V, Konongoi SL, et al. Genomic surveillance of Rift Valley fever virus: from sequencing to lineage assignment. *BMC Genomics* 2022; **23**: 520.
- 22 Soumaré POL, Freire CCM, Faye O, et al. Phylogeography of Rift valley fever virus in Africa reveals multiple introductions in Senegal and Mauritania. *PLoS One* 2012; **7**: e35216.
- 23 Samy AM, Peterson AT, Hall M. Phylogeography of Rift Valley fever virus in Africa and the Arabian Peninsula. *PLoS Negl Trop Dis* 2017; **11**: e0005226.
- 24 Frankema E, Green E, Hillbom E. Success and failure of European settler farming in colonial Africa. 2014. <https://www.aehnetwork.org/working-papers/success-and-failure-of-european-settler-farming-in-colonial-africa/> (accessed March 11, 2024).
- 25 Weaver SC. Evolutionary influences in arboviral disease. *Curr Top Microbiol Immunol* 2006; **299**: 285–314.
- 26 Gachohi JM, Njenga MK, Kitala P, Bett B. Modelling vaccination strategies against Rift Valley fever in livestock in Kenya. *PLoS Negl Trop Dis* 2016; **10**: e0005049.
- 27 Turner MD, Schlecht E. Livestock mobility in sub-Saharan Africa: a critical review. *Pastoralism* 2019; **9**: 13.