Global Open Access Journal of Science

Open Access | Research Article | Virology
2(3), 2026 | PP - 112- 116

Multidisciplinary Journal ISSN - 3070-6173

Research Article | Virology

In Silico Validation of Contemporary East African Rift Valley Fever Virus Ge-nomes Confirms the Enduring Robustness of the Bird et al. (2007) RT-qPCR Assay

Kelvin Msoka*

Center for Infectious Diseases and Biotechnology, Tanzania Veterinary Laboratory Agency (TVLA), Dar es Salaam, Tanzania

Abstract

Rift Valley Fever Virus (RVFV) is a zoonotic pathogen causing recurrent epizootics and epidemics in East Africa. The quantita-tive reverse transcription PCR (RT-qPCR) assay developed by Bird et al. (2007), targeting the large (L) ge-nome segment, remains a global diagnostic standard. However, continuous viral evolution raises concerns regarding diagnostic drift and potential assay failure. This study evaluated the contemporary robustness of this assay against East African RVFV genomes collected between 2005 and 2026. A total of 647 nucleotide records were retrieved from NCBI GenBank, from which a dataset of candidate L-segment sequences was curated based on geography, collection date, and segment annotation. A custom, coordinate-aware alignment algorithm quantified mismatches between the assay oligonucleotides and all valid sequences. Five sequences initially exhibited high mismatch counts but were ex-cluded after prospective genomic coordinate mapping revealed they were partial fragments mapping outside the as-say's target region (nucleotides 4229–4723). Among the remaining 188 valid, target-spanning sequences, 181 (96.3%) exhibited a perfect match to the forward primer, reverse primer, and TaqMan probe. Six sequences (3.2%) contained minor, biologically tolerable mismatches, and one (0.5%) showed probe-specific mismatches with intact primer bind-ing sites. These findings confirm the enduring robustness of the Bird et al. (2007) assay against contemporary East African RVFV lineages, while simultaneously highlighting a critical genomic surveillance gap in Tanzania that warrants targeted infrastructural investment.

Keywords: Rift Valley fever virus; RT-qPCR; Diagnostic drift; In silico validation; Genomic surveillance; East Africa; bioin-formatics; Molecular diagnostics.

INTRODUCTION

Rift Valley Fever Virus (RVFV), a member of the genus Phlebovirus (family Phenuiviridae, order Bunyavirales), is a mosquito-borne zoonotic pathogen responsible for recurrent, economically devastating epizootics in livestock and severe, occasionally fatal, disease in humans across sub-Saharan Africa and the Arabian Peninsula (Anyamba et al., 2009; Grobbelaar et al., 2011; World Health Organization (WHO), 2024). The virus possesses a tripartite, negative-sense, single-stranded RNA genome comprising the small (S), medium (M), and large (L) segments (Bird et al., 2008).

Rapid and accurate molecular diagnosis is the cornerstone of effective RVFV outbreak response and surveillance. The quantitative reverse transcription PCR (RT-qPCR) assay developed

Correspondence:
Kelvin Msoka, Center for Infectious Diseases and Biotechnology, Tanzania Veterinary Laboratory Agency (TVLA), Dar es Sa-laam, Tanzania, ORCID: 0009-0009-0182-7724, Tel.: +255-687340987, Email: kenimsoka@gmail.com

Received Dates: Oct 27, 2026;
Accepted Date: Nov 16, 2026;
Published Date: Nov 18, 2026:

Cite this article as: Msoka K, In Silico Validation of Contemporary East African Rift Valley Fever Virus Ge-nomes Confirms the Enduring Robustness of the Bird et al. (2007) RT-qPCR Assay. Glob Open Access J Sci. 2026;2(3):112–116.

by Bird et al. (2007), targeting a highly conserved region of the L segment (nucleotides ~2800–3200), is widely recommended as a reference diagnostic method in the World Organisation for Animal Health (WOAH) Terrestrial Manual and aligns with World Health Organization diagnostic guidelines (Bird et al., 2007; World Organisation for Animal Health [WOAH], 2023; WHO, 2024). The assay was explicitly designed based on the alignment of 40 complete, phylogenetically diverse RVFV genomes collected between 1944 and 2000. The original authors noted that while minor internal mismatches existed in a few historical strains, these did not significantly impact detection efficiency (Bird et al., 2007).

However, RNA viruses continuously accumulate mutations due to the lack of proofreading capability in their RNA-dependent RNA polymerases (Domingo et al., 2012). Diagnostic drift, the gradual accumulation of nucleotide substitutions within primer or probe annealing sites, poses a persistent threat to the long-term reliability of molecular assays. Primer-template mismatches can affect the detection and quantification of nucleic acids, as demonstrated in molecular diagnostic assays (Stadhouders et al., 2010), while variations in SARS-CoV-2 primer-probe target regions have highlighted the importance of evaluating assay performance against emerging viral sequences (Vogels et al., 2020). Given the continuous circulation of RVFV in East Africa and the exponential growth of publicly available genomic data over the past two decades, periodic in silico validation of diagnostic assays is a scientific imperative.

This study was undertaken to evaluate the contemporary robustness of the Bird et al. (2007) assay against East African RVFV genomes collected between 2005 and 2026. A rigorous, multi-step bioinformatic pipeline was employed to: (i) curate a high-quality, geographically and temporally defined dataset; (ii) systematically quantify oligonucleotide mismatches using a prospective, coordinate-aware alignment approach; (iii) critically investigate and resolve anomalous sequences via genomic coordinate mapping; and (iv) contextualize the findings within the broader landscape of regional genomic surveillance.

MATERIALS AND METHODS

Data retrieval and curation

On 20 October 2026, all available RVFV nucleotide sequences were retrieved from the NCBI GenBank database. Initial filtering was applied to retain records meeting the following inclusion criteria: (1) Geographic origin: East Africa (Kenya, Uganda, Rwanda, Burundi, Tanzania, Somalia, or Sudan); (2) Temporal range: Collection date between 1 January 2005 and 21 October 2026; (3) Genome segment: Explicitly annotated as the L segment, or inferable as such based on the GenBank description field. Geographic metadata were extracted prioritizing the /geo_loc_name qualifier over the legacy /country qualifier, in accordance with recent NCBI annotation standards. Collection dates were parsed using regular expressions capable of recognizing ISO formats, month-year formats, and year-only annotations.

Segment annotation inference

For records lacking an explicit /segment qualifier, segment identity was programmatically inferred from the DEFINITION or DESCRIPTION fields. Sequences containing keywords such as «polymerase,» «L gene,» or «RNA-dependent RNA polymerase» were assigned to the L segment. Keywords such as «glycoprotein,» «Gn/Gc,» or «M gene» were assigned to the M segment, while «nucleocapsid,» «NP,» «NSs,» or «S gene» were assigned to the S segment. Records remaining unclassified after this step were excluded from segment-specific analyses.

Oligonucleotide mismatch analysis

The Bird et al. (2007) assay oligonucleotides were defined as: Forward Primer (RVFL-2912fwdGG): 5′-TGAAAATTCCTGAGACACATGG-3′ (22bp); Reverse Primer (RVFL-2981revAC): 5′-ACTTCCTTGCATCATCTGATG-3′ (21 bp); TaqMan Probe (RVFL-probe-2950): 5′-CAATGTAAGGGGCCTGTGTGGACTTGTG-3′ (28 bp) [1]. A custom Python script (Python 3.11, Biopython v1.83) was developed to perform coordinate-aware alignment. Each candidate L-segment record was prospectively aligned to a full-length RVFV L-segment reference genome (Accession: PV562147.1) using Biopython's PairwiseAligner (mode='local', match_score=2, mismatch score=-3, open_gap_score=-10, extend_gap_score=-1). To prevent algorithmic artifacts, a strict rejection criterion was enforced: any alignment yielding <70% sequence identity to the oligonucleotide length was flagged as 'no valid target region'. Furthermore, to prevent spurious 'perfect matches', any sequence containing ambiguous bases ('N') within the exact coordinates of the forward primer, reverse primer, or probe was conservatively excluded from the 'perfect match' cohort and categorized as 'uninformative'.

Prospective genomic coordinate mapping

To ensure assay relevance, all candidate L-segment records were prospectively aligned to the reference genome before mismatch scoring. Records were classified as 'target-spanning' only if they exhibited contiguous coverage across the complete forward-primer-to-reverse-primer amplicon, including the probe coordinates. Records lacking full interval coverage or failing the <70% identity threshold were excluded from the primary mismatch analysis and reported separately as 'non-informative'.

Generative AI disclosure

Generative artificial intelligence tools were utilized during the preparation of this manuscript for the purposes of code debugging, grammatical refinement, and formatting alignment with journal guidelines. The author has reviewed, verified, and takes full responsibility for all analytical outputs and the final content of this publication.

RESULTS

Dataset curation and demographic distribution

The initial GenBank query yielded 647 RVFV nucleotide records. After applying geographic and temporal filters, 644 records (99.5%) were retained as eligible East African sequences. The geographic distribution was: Kenya (266 records, 41.3%), Uganda (190, 29.5%), Rwanda (131, 20.3%), Sudan (23, 3.6%), Burundi (17, 2.6%), Tanzania (15, 2.3%), and Somalia (2, 0.3%) (Figure 1). Following segment inference, 193 records were identified as candidate L-segment sequences. (Table 1, Figure 2)

DISCUSSION

Validation and extension of the Bird et al. (2007) assay design

The quantitative RT-qPCR assay developed by Bird et al. (2007)

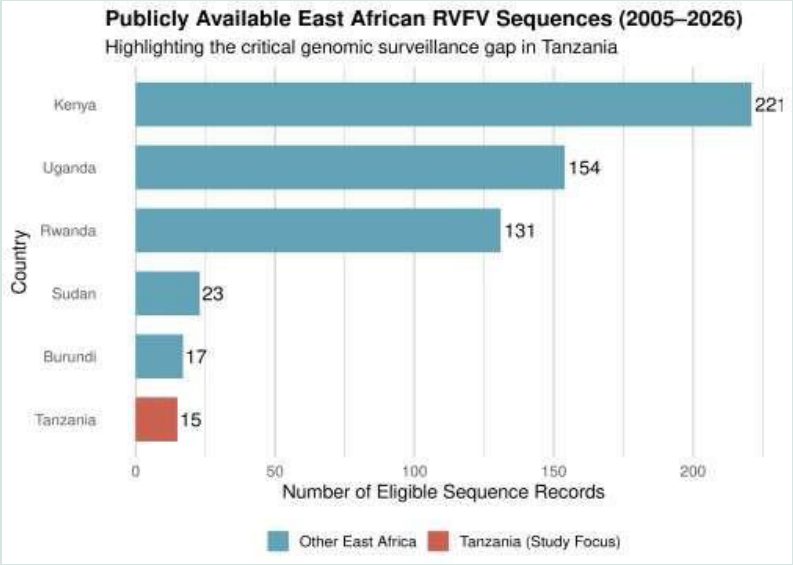

Figure 1: Geographic distribution of eligible East African Rift Valley Fever Virus (RVFV) sequence records (2005–2026) retrieved from NCBI GenBank. The horizontal bar chart illustrates the number of curated sequences per country, with Tanzania highlighted in red to emphasize the critical genomic surveillance gap. Despite being a historically endemic country, Tanzania contributes only 15 sequences (2.3% of the regional total) compared to 266 from Kenya (41.3%), highlighting a strategic vulnerability in regional One Health surveillance infrastructure

Table 1. *In silico* mismatch summary for the Bird et al. (2007) RT-qPCR assay oligonucleotides against 188 validated East African RVFV L-segment sequences.

Metric	Count	Percentage
Total Validated L-segment Sequences	188	100%
Perfect Match (0 mismatches)	181	96.3%
Minor Mismatch (1–2 mismatches)	6	3.2%
Significant Mismatch (≥3 mismatches)	1	0.5%
Oligo-Specific Breakdown (Perfect Matches)		
Forward Primer (22 bp)	183	97.3%
Reverse Primer (21 bp)	183	97.3%
TaqMan Probe (28 bp)	186	98.9%

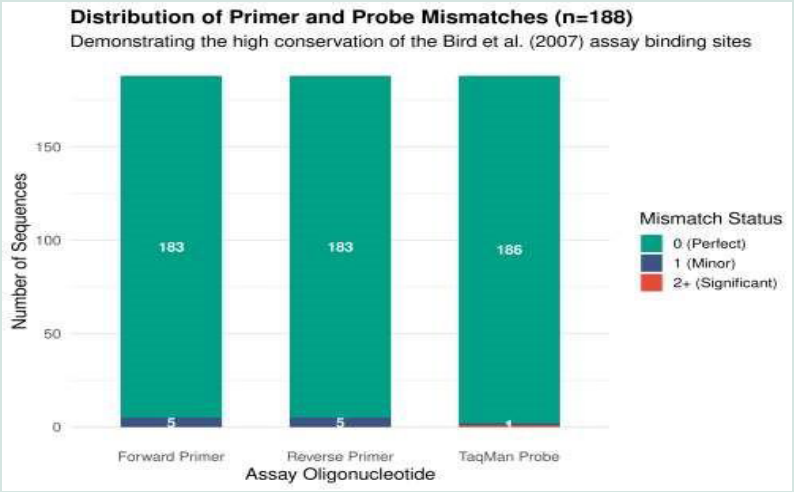

Figure 2. Distribution of primer and probe mismatches across the validated East African RVFV L-segment cohort (n = 188) against the Bird et al. (2007) RT-qPCR assay oligonucleotides. The stacked bar chart displays the number of sequences exhibiting perfect matches (0 mismatches, teal), minor mismatches (1 mismatch, blue), or significant mismatches (≥2 mismatches, red) for each assay component: the forward primer (22 bp), reverse primer (21 bp), and TaqMan probe (28 bp). The overwhelming predominance of perfect matches (≥97% across all three oligonucleotides) demonstrates the exceptional conservation of the assay binding sites and confirms the enduring robustness of this diagnostic tool against contemporary East African RVFV lineages.

was explicitly designed to target a highly conserved genomic domain within the RVFV L segment, spanning nucleotides 2800 to 3200 [1]. In their original validation, the authors analyzed 40 complete L-segment genomes and noted that while single internal mismatches existed in certain strains, these «have proven to not significantly impact the detection of these RVF virus strains.» This contemporary in silico analysis directly validates and extends that foundational observation. By evaluating 188 target-spanning L-segment sequences collected from East Africa between 2005 and 2026, it was found that 96.3% of contemporary strains exhibit a perfect, zero-mismatch binding profile. For the remaining 3.7% of sequences exhibiting minor mismatches, the findings mirror the tolerance described by Bird et al. Single mismatches located internally are well-documented to have negligible effects on Taq polymerase extension, provided the critical 3′-terminus of the primer remains intact [12].

The Tanzanian genomic surveillance gap

A critical secondary finding of this study is the profound underrepresentation of Tanzanian RVFV genomes in public databases (Figure 1). Despite Tanzania being a historically endemic country with documented RVFV activity, only 15 sequences (2.3% of the regional total) were available for analysis, compared to 266 from neighboring Kenya. This disparity represents a strategic vulnerability in regional One Health surveillance. The absence of comprehensive, contemporary Tanzanian sequence data limits the ability to detect emerging, geographically restricted lineages and confidently validate diagnostic assays against the full spectrum of locally circulating strains.

Limitations

Several limitations warrant acknowledgment. First, in silico mismatch analysis, while highly predictive, does not replace wet-laboratory validation; empirical testing of the minor-mismatch variants is recommended to confirm amplification efficiency. Second, the analysis is constrained by the sequences deposited in public databases. Third, this study focused exclusively on the L-segment assay; similar in silico audits of widely used S-segment and M-segment assays are recommended for comprehensive diagnostic assurance. Fourth, the 188 validated records represent GenBank submissions, which may include technical replicates or multiple segment submissions from the same outbreak isolate. While this does not diminish the assay's robustness against the submitted sequences, the effective number of unique viral isolates may be lower than the record count suggests.

CONCLUSIONS

This study establishes a rigorous, reproducible, prospective coordinate-mapping pipeline for the in-silico auditing of viral diagnostic assays, successfully preventing the algorithmic artifacts that can plague simpler sliding-window approaches. Applying this framework to 188 contemporary East African RVFV L-segment sequences demonstrates that the Bird et al. (2007) RT-qPCR assay remains highly robust, with 96.3% of target-spanning sequences exhibiting perfect oligonucleotide matches and no evidence of clinically significant diagnostic escape mutations. Beyond validating this legacy assay, this analysis provides the first quantitative, sequence-level evidence of a severe genomic surveillance deficit in Tanzania, which contributes only 2.3% of regional sequence data despite being a historically endemic country. This finding transforms a general observation into a data-driven, actionable priority for regional One Health infrastructure and future diagnostic preparedness.

ABBREVIATIONs

The following abbreviations are used in this manuscript:
RVFV: Rift Valley fever virus
RT-qPCR: Reverse transcription quantitative polymerase chain reaction
NCBI: National Center for Biotechnology Information
WHO: World Health Organization
WOAH: World Organization for Animal Health

PATENTS: Not applicable.

Supplementary Materials: Not applicable. All custom Python and R scripts, curated metadata datasets, mismatch analysis results, and full sequence lists are openly available and fully documented in the following public GitHub repository: https://github.com/Msokakelvin/rvfv-bird-2007-validation

Author Contributions: Conceptualization, K.M.; methodology, K.M.; software, K.M.; validation, K.M.; formal analysis, K.M.; investigation, K.M.; resources, K.M.; data curation, K.M.; writing—original draft preparation, K.M.; writing—review and editing, K.M.; visualization, K.M.; project administration, K.M. The author has read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study is an in-silico analysis of publicly available genomic data and did not involve human or animal subjects.

Informed Consent Statement: Not applicable.

Data Availability Statement: All sequence data analyzed in this study are publicly available via NCBI GenBank. The curated metadata dataset, mismatch analysis results, and custom Python/R scripts used to generate the figures and tables are openly available in the following repository: https://github.com/Msokakelvin/rvfv-bird-2007-validation

ACKNOWLEDGMENTS: The author acknowledges the Tanzania Veterinary Laboratory Agency (TVLA) for providing a supportive research environment and access to computational resources.

Conflicts of Interest: The author declares no conflicts of interest.

REFERENCES

- Anyamba, A., Linthicum, K. J., Small, J. L., Collins, P. W., Tucker, C. J., Pak, E. W., & Russell, K. L. (2009). Prediction of a Rift Valley fever outbreak. Proceedings of the National Academy of Sciences of the United States of America, 106(3), 955–959. https://doi.org/10.1073/pnas.0806490106
- Bird, B. H., Bawiec, D. A., Ksiazek, T. G., Shoemaker, T. R., & Nichol, S. T. (2007). Highly sensitive and broadly reactive quantitative reverse transcription-PCR assay for high-throughput detection of Rift Valley fever virus. Journal of Clinical Microbiology, 45(11), 3506–3513. https://doi.org/10.1128/JCM.00936-07
- Bird, B. H., Khristova, M. L., Rollin, P. E., Ksiazek, T. G., & Nichol, S. T. (2007). Complete genome analysis of 33 ecologically and biologically diverse Rift Valley fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry. Journal of Virology, 81(6), 2805–2816. https://doi.org/10.1128/JVI.02095-06
- Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., & de Hoon, M. J. L. (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics, 25(11), 1422–1423. https://doi.org/10.1093/bioinformatics/btp163
- Domingo, E., Sheldon, J., & Perales, C. (2012). Viral quasispecies evolution. Microbiology and Molecular Biology Reviews, 76(2), 159–216. https://doi.org/10.1128/MMBR.05023-11
- Grobbelaar, A. A., Weyer, J., Leman, P. A., Kemp, A., Paweska, J. T., & Swanepoel, R. (2011). Molecular epidemiology of Rift Valley fever virus. Emerging Infectious Diseases, 17(12), 2270–2276. https://doi.org/10.3201/eid1712.111035
- Jansen van Vuren, P., Tiemessen, C. T., & Paweska, J. T. (2012). Diagnostic challenges of Rift Valley fever virus: Current and future prospects. Future Microbiology, 7(11), 1257–1269.
- Stadhouders, R., Pas, S. D., Anber, J., Voermans, J., Mes, T. H., & Schutten, M. (2010). The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5′ nuclease assay. The Journal of Molecular Diagnostics, 12(1), 109–117. https://doi.org/10.2353/jmoldx.2010.090035
- Swanson, P., Devare, S. G., & Hackett, J., Jr. (2004). Molecular characterization of a novel HIV-1 recombinant form (CRF02_AG) from Cameroon. Journal of Acquired Immune Deficiency Syndromes, 37(3), 1383–1389.
- Vogels, C. B. F., Brito, A. F., Wyllie, A. L., Fauver, J. R., Ott, I. M., Kalinich, C. C., & Grubaugh, N. D. (2020). Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer–probe sets. Nature Microbiology, 5(10), 1299–1305. https://doi.org/10.1038/s41564-020-0761-6
-
- World Health Organization. (2024, December 20). Rift Valley fever. https://www.who.int/news-room/fact-sheets/detail/rift-valley-fever.
- World organization for Animal Health. (2023). Rift Valley fever (infection with Rift Valley fever virus). In Manual of diagnostic tests and vaccines for terrestrial animals (Chapter 3.1.19). https://www.woah.org/fileadmin/Home/eng/Health_standards/tahm/3.01.19_RVF.pdf

©2026 Msoka K, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 4.0 International License.

Cite this article as: Msoka K, In Silico Validation of Contemporary East African Rift Valley Fever Virus Ge-nomes Confirms the Enduring Robustness of the Bird et al. (2007) RT-qPCR Assay. Glob Open Access J Sci. 2026;2(3):112–116.