JIDC | THE JOURNAL OF INFECTION IN DEVELOPING COUNTRIES

Emerging Problems in Infectious Diseases

Clock-like behavior across *Orthoebolavirus* species suggests alternative evolution of the 2026 Bundibugyo virus outbreak

Gustavo Sganzerla Martinez[1,2,3 #], Anuj Kumar[3 #], Mansi Dutt[3], Misaki Wayengera[4,5,6], Wilber Sabiiti[4,7], David J Kelvin[1,3], Henry Kyobe Bosa[4,5,8]

[1] *Basic Medical Sciences, Shantou University Medical College, Shantou, Guangdong, People's Republic of China*
[2] *Immunology, Shantou University Medical College, Shantou, Guangdong, People's Republic of China*
[3] *Microbiology and Immunology, Dalhousie University, Faculty of Medicine, Halifax, Nova Scotia, B3H4H7, Canada*
[4] *Interdisciplinary Consortium for Epidemic Research and Response (ICER), Kampala, Uganda*
[5] *Ministry of Health, Kampala, Uganda*
[6] *Departments of Immunology, Molecular Biology and Pathology, School of Biomedical Sciences, College of Health Sciences, Makerere University, Kampala, Uganda*
[7] *Division of Infection and Global Health, School of Medicine, University of St Andrews, Scotland, UK*
[8] *Rear Public Health Laboratory, Kampala, Uganda*

# Authors contributed equally and share first authorship of this work.

Abstract

Introduction: Recent increase in reported cases of Bundibugyo virus (species *Orthoebolavirus bundibugyoense*) in the Democratic Republic of the Congo and Uganda highlight the importance of understanding their evolutionary dynamics.

Methodology: We analyzed publicly available genomes obtained from the Pathoplexus database representing 3 medically relevant species of the *Orthoebolavirus* genus: Ebola virus (n = 3,388), Zaire virus (n = 166), and Bundibugyo virus (n = 49). We investigated whether genetic divergence is associated with sampling time utilizing root-to-tip regression analyses derived from temporally calibrated phylogenies as an indicator of molecular clock behavior.

Results: Sudan virus ($R^2 = 0.855$) and Ebola virus ($R^2 = 0.709$) exhibited significant temporal structures between sampling time and genetic divergence. Bundibugyo virus displayed strong clock-like behavior after exclusion of 2026 genomes ($R^2 = 0.983$), which showed a marked substitution deficit relative to the historical molecular clock and deviated substantially from the inferred regression. Date randomization tests supported the presence of temporal signal across all datasets ($p < 0.001$), indicating that the temporal signals exceeded expectations under a null model without temporal structure.

Conclusions: These results provide a comparative assessment of the temporal structures across 3 medically relevant *Orthoebolavirus* species. While historical genomes largely conformed to molecular clock expectations, the inclusion of genomes from the ongoing 2026 Bundibugyo virus outbreak deviated from these projections, suggesting an alternative evolutionary process for 2026 Bundibugyo virus, such as viral persistence in immune-privileged anatomical sites of previously infected individuals or an independent zoonotic spillover event in an alternative species, in which the virus evolved within its natural reservoir prior to human infection.

Key words: *Orthoebolavirus*; Ebola; evolution.

J Infect Dev Ctries 2026; 20(6):767-771. doi:10.3855/jidc.23736

(Received 19 June 2026 – Accepted 23 June 2026)

Introduction

The *Orthoebolavirus* genus (*Filoviridae* family) comprises enveloped single-stranded RNA viruses that can infect humans and cause hemorrhagic fever. Four viruses in the *Orthoebolavirus* genus can cause diseases in humans: (1) Ebola virus (species *Orthoebolavirus zairense*), (2) Sudan virus (species *Orthoebolavirus sudanense*), (3) Bundibugyo virus (species *Orthoebolavirus bundibugyoense*), and (4) Taï Forest virus (species *Orthoebolavirus taiense*) [1–3]. The largest *Orthoebolavirus* outbreak was caused by *Orthoebolavirus zairense* in 2014–2016 in West Africa and resulted in more than 28,600 infected individuals with 11,325 recorded deaths [4,5]. Moreover, the World Health Organization (WHO) declared a public health emergency of international concern on 17 May 2026 following an outbreak of the *Orthoebolavirus bundibugyoense* in the eastern region of the Democratic Republic of the Congo (DRC) and Uganda. As of 5 June 2026, DRC and Uganda have recorded 471 confirmed cases with 84 confirmed deaths, totaling a case fatality rate of 18% [6].

Genomic surveillance is an important component of disease outbreak research as it enables the reconstruction of evolutionary relationships and transmission history through phylogenetic analyses.

Temporal phylogenetic approaches rely on measurable temporal structures where genetic divergence can be explained through sampling time. Evaluating whether sampled viral genomes evolve under clock-like assumptions is therefore a requisite for downstream evolutionary and phylodynamic investigations. Here, the temporal structure of 3 datasets comprising medically relevant *Orthoebolavirus* species (Ebola virus, Sudan virus, and Bundibugyo virus) was assessed by examining the relationship between sampling time and evolutionary divergence using root-to-tip (RTT) regression analyses derived from temporally calibrated phylogenetic trees. The aim was to evaluate the presence of clock-like behavior and the suitability of these datasets for temporal evolutionary analyses.

Methodology

The Pathoplexus database was searched on 2 July 2026 and a total of 166 Sudan virus genomes, [7] 3,388 Ebola virus genomes [8], and 49 Bundibugyo virus genomes [9] was obtained. These numbers represent genomes filtered based on the presence of collection date and genome length of ≥ 18,000 base pairs. Fifteen Bundibugyo virus genomes represent isolates from Uganda in the 2026 outbreak [10–12]. All analyses were conducted in each *Orthoebolavirus* species. Genomic sequences were aligned using MAFFT (version 7.5.2) [13] in fast alignment mode. The resulting alignments were used to infer maximum likelihood phylogenetic trees with IQ-TREE2 (version 2.3.6) [14] under the GTR + G substitution model with 1,000 bootstrap replicates.

Recombination was assessed using Gubbins (version 3.0) [15]. Analyses were performed over 5 iterative cycles using FastTree for rapid tree reconstruction. Recombination blocks were identified using a minimum cluster of 5 single nucleotide polymorphisms (SNPs), and recombination signals were summarized using the number of inferred recombination events and recombination-to-mutation ratio (r/m).

Temporal structure was assessed using a RTT regression framework. RTT distances were estimated using TreeTime (version 0.9.4) [16] and used as input for downstream statistical analyses. Phylogenetic trees were rooted using a least-square approach. Temporal signal was evaluated by regressing TreeTime-derived RTT distances against sampling dates using ordinary least squares (OLS), with the coefficient of determination (R^2) used as a measurement of clock-like behavior. To evaluate robustness to outliers, terminal taxa with RTT distances exceeding ± 3 standard deviations from the mean were flagged and excluded in a sensitivity analysis, and regressions were recomputed. Temporal signal significance was evaluated using a date randomization test (DRT). Sampling dates were randomly permuted across taxa while preserving tree topology and RTT structure. A total of 1,000 permutations were performed. For each permuted dataset, OLS regression was repeated and R^2 values were recorded to generate a null distribution. Empirical significance was calculated as the proportion of permuted datasets with R^2 values greater than or equal to the observed value. The 95th percentile of the null distribution was used as an empirical threshold for non-temporal signal. As RTT regression depends on tree topology and rooting, the results were interpreted as evidence of temporal signal consistent with molecular clock expectation rather than formal confirmation of strict clock-like evolution. The inclusion of a 2026 genome in the Bundibugyo virus dataset rendered a deviation from the expected temporal trend observed in earlier sampling years (2007–2012). To quantify this deviation, a regression model was fitted using only the historical dataset (2007–2012) and extrapolated to 2026 under a constant rate to provide an exploratory reference trajectory. Residual variance of the baseline regression was estimated as:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2}$$

Where y_i is the observed RTT and $\hat{y}_i$ is the value predicted by the regression model. For the 2026 genome, the evolutionary deviation from expectation was quantified as:

$$\Delta Distance_j = y_j - \hat{y}_j$$

To assess whether this deviation exceeded expectations under the fitted model, a studentized prediction residual was computed:

$$t_j = \frac{y_j - \hat{y}_j}{\sigma\sqrt{1 + \frac{1}{n} + \frac{(x_j - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}}$$

Where x_j is the sampling date of the 2026 genome and x_j is the mean sampling date of the baseline dataset. Under OLS assumptions, this statistic follows an approximated Student's *t*-distribution with *n* - 2 degrees of freedom.

Multiple sequence alignment, phylogenetic tree building, and clock estimations were executed in the High-Performance Computing (HPC) cluster Nibi, hosted by the Digital Research Alliance of Canada. TreeTime and Gubbins were executed via their

containerized Docker images (https://hub.docker.com/r/snads/treetime and https://hub.docker.com/r/sangerpathogens/gubbins, respectively) via Apptainer (version 1.4.5) to maintain compatibility with the HPC environment. Downstream regression analyses and plotting were performed locally with Python (version 3.12) scripts.

Data availability

The *Orthoebolavirus* genomes considered for this analysis are all public, non-embargoed data deposited to the Pathoplexus database. DOIs were generated for each *Orthoebolavirus* species with data analyzed: Sudan virus (10.62599/PP_SS_2245.1) [7], Ebola virus (10.62599/PP_SS_2246.1) [8], and Bundibugyo virus (10.62599/PP_SS_2212) [9]. We acknowledge the involvement of the original data producers and their respective laboratories for their efforts towards the broader infectious disease genomics community.

Results

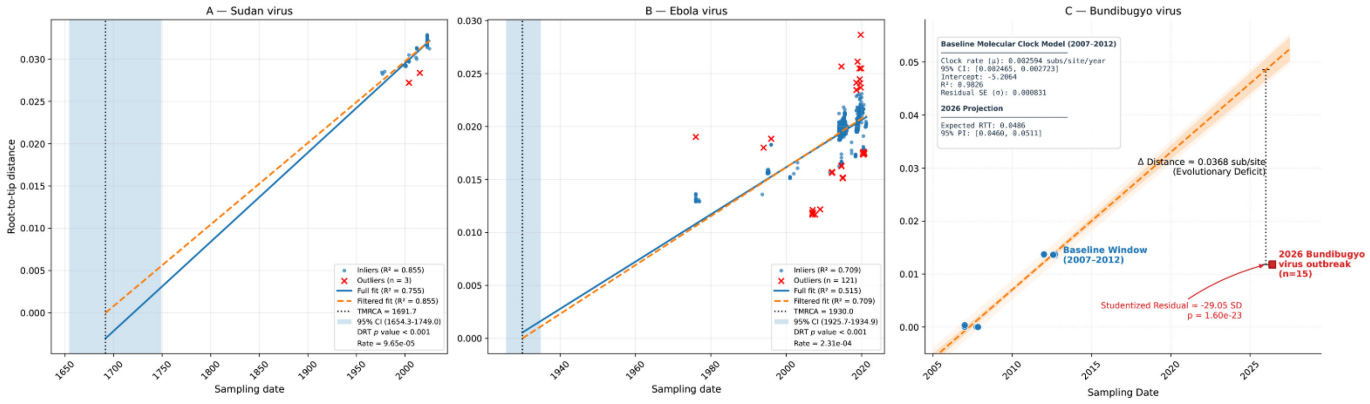
RTT regression analyses revealed heterogenous but consistently significant temporal structures across the 3 evaluated *Orthoebolavirus* species (Figure 1). All datasets showed a positive association between sampling dates and genetic divergence, consistent with measurable molecular clock signal.

Sudan virus (*Orthoebolavirus sudanense*) (Figure 1A) displayed a temporal signal $R^2 = 0.755$ on the full dataset. After removing outliers (n = 3), the temporal signal increased to $R^2 = 0.855$ indicating that the data is compatible with clock-like evolutionary pattern. The time to most recent common ancestor (TMRCA) was estimated in approximately 1,692 (95% CI: 1652.7–1748.9) with a rate of 9.65e-05 substitutions per site per year.

Ebola virus (*Orthoebolavirus zairense*) (Figure 1B) had an initial weaker temporal signal $R^2 = 0.515$. A total of 121 outliers were identified and after their removal a stronger temporal signal $R^2 = 0.709$ was obtained. TMRCA on the sampled data was identified as 1930 (95% CI: 1925.5–1934.7) with a rate of $2.31e^{-04}$ substitutions per site per year.

To investigate the evolutionary pace of sequences from the 2026 active Bundibugyo virus outbreak, first an RTT regression analysis was performed across all sampled Bundibugyo virus (*Orthoebolavirus bundibugyoense*) and a reduced temporal signal and poor model fit ($R^2 = 0.18$) was obtained. This distortion suggested potential violations of a strict molecular clock assumption when the full dataset was analyzed. To formally evaluate temporal structure without bias from the 2026 genomes, an out-of-sample linear prediction framework was implemented (Figure 1C). A baseline model was fitted exclusively to sequences from 2007 and 2012, which exhibited a strong clock-

Figure 1. Root-to-tip (RTT) regression analyses of *Orthoebolavirus* genomes.



RTT distances were extracted from temporally calibrated phylogenetic trees using TreeTime (version 0.9.4) under a strict molecular clock assumption. Sampling dates were regressed against genetic divergence of Ebola for 3 *Orthoebolavirus* species and individually analyzed. **A.** Sudan virus; **B.** Ebola virus; **C.** Bundibugyo virus.

Panels A and B show the full dataset regression alongside an additional regression with outlier genomes with RTT distances exceeding ± 3 standard deviations from the mean RTT excluded. The strength of the temporal signal was assessed as the coefficient of determination R^2 and supported by date randomization tests (DRT) with 1,000 permutations. Time to most recent common ancestor (TMRCA) is presented as a fixed date of the regression analysis as well as its 95% confidence internal (CI). For the Bundibugyo virus dataset (**panel C**), first, the genomes from 2007 and 2012 (blue dots) were regressed, modeling RTT distance as a function of sampling date. Under the 2007–2012 constant rate assumption, a projected trajectory was extended to 2026. Genomes from Uganda comprising the 2026 active outbreak (n = 15) sit below the projected trajectory. The vertical dotted line indicates the difference between observed (0.011793) and expected (0.048551) RTT distances. A test encompassing Studentized prediction residuals was conducted to assess whether the RTTs of the 2026 Bundibugyo genomes were consistent with a substantial deviation from the historical molecular clock inferred from the 2007–2012 genomes. The reported rates are RTT regression estimates and do not represent full model-based substitution rates.

like behavior ($R^2 = 0.98$) and an estimated substitution rate of 2.25×10^{-3} substitutions per site per year (95% CI $2.24 \times 10^{-3} – 2.27 \times 10^{-3}$). Projection of the baseline molecular clock model to 2026 yielded an expected RTT distance of approximately 0.0486 substitutions per site (95% CI: 0.0460 – 0.0511). In contrast, the 2026 sequences (n = 15) displayed substantially lower divergence value corresponding to an evolutionary deficit of 0.0368 substitutions per site (observed 0.011793; expected 0.48551). Standardized prediction residuals indicated that the 2026 genomes represent extreme outliers relative to the baseline evolutionary model, approaching 30 standard deviations below expectation. These results indicate that the sequences from the 2026 ongoing outbreak substantially deviate from the temporal expectations of historical Bundibugyo virus baseline of the historical clock model, suggesting a marked disruption in the clock-like accumulations over the 2007–2026 timeframe.

Across the 3 species, a date randomization test (DRT) consistently supported temporal structures ($p < 0.001$) indicating the observed associations between sampling date and genetic divergence exceeded expectations under a null model with no temporal signal. Measurable clock-like behavior in all 3 tested *Orthoebolavirus* species was identified. No evidence of blocks derived from recombination events was identified in any of the queried species.

Discussion

The analyses demonstrated that 3 medically relevant *Orthoebolavirus* species contained measurable and statistically supported temporal structures. RTT regression analyses based on TreeTime-derived phylogenetic distances revealed a consistent positive association between sampling time and genetic divergence for both Ebola virus and Sudan Virus datasets. Bundibugyo virus showed a strong temporal signal only when the 2026 genomes were excluded from the analysis. RTT regression was used here as an exploratory framework to assess the presence and consistency of temporal signal rather than as a formal phylodynamic inference method, making the results indicators of compatibility with molecular clock expectations, instead of direct estimates of evolutionary dynamics.

Conclusions

Two non-mutually exclusive hypotheses were identified that may explain the observed deviation of the 2026 Bundibugyo virus genomes from the historical molecular clock. First, the reduced accumulation of substitutions over a ~14-year period may be consistent with viral persistence in immune-privileged anatomical sites of previously infected individuals [17,18]. Under this scenario, prolonged or intermittent replication within a long-term host reservoir could result in deviations from the substitution rates inferred under continuous transmission dynamics. Second, the observed pattern may reflect an independent zoonotic spillover event in alternative species, in which the virus evolved within its natural reservoir prior to human infection [19,20], leading to a distinct evolutionary trajectory resulting in distinct substitution rates and phylogenetic trajectories when introduced to humans. Discriminating between these two hypotheses will require integration of genomic, epidemiological, and ecological data, including investigation of links to previously infected individuals and expanded sampling of potential wildlife reservoirs and intermediate hosts to determine whether the 2026 lineage reflects cryptic circulation of a separate emergence event.

Acknowledgements

This study was supported by awards from the Canada Research Chair program, Moderna Global Fellowship 2024 (2024-MGF-000316 [91353]), and the Li-Ka Shing Foundation. AK is a Moderna Global Fellow and DJK is the Canada Research Chair in Translational Vaccinology and Inflammation. Authors from Uganda were supported by the Government of Uganda.

Corresponding author

David J Kelvin, PhD.
5850 College Street, Halifax, NS, Canada B3H4H7.
Tel: +1 (647) 529-3556
Email: david.kelvin@dal.ca

Conflict of interest

AK, GSM, MD, and DJK are shareholders of the company BioForge Canada Limited. BioForge Canada Limited is a company that uses bioinformatics in immunological approaches for the monitoring, prevention, and treatment of infectious diseases. The authors disclose that the interests of BioForge Canada Limited had no impact in this study. DJK is a scientific advisory board member for Emergent BioSolutions.

References

1. Hasan S, Ahmad S, Masood R, Saeed S (2019) Ebola virus: a global public health menace: a narrative review. J Family Med Prim Care 8: 2189–2201. doi: 10.4103/jfmpc.jfmpc_297_19.
2. Malik S, Kishore S, Nag S, Dhasmana A, Preetam S, Mitra O, León-Figueroa DA, Mohanty A, Chattu VK, Assefi M, Padhi BK, Sah R (2023) Ebola virus disease vaccines: development, current perspectives & challenges. Vaccines 11: 268. doi: 10.3390/vaccines11020268.

3. Rojas M, Monsalve DM, Pacheco Y, Acosta-Ampudia Y, Ramírez-Santana C, Ansari AA, Gershwin ME, Anaya JM (2020). Ebola virus disease: an emerging and re-emerging viral threat. J Autoimmun 106: 102375. doi: 10.1016/j.jaut.2019.102375.
4. Coltart CEM, Lindsey B, Ghinai I, Johnson AM, Heymann DL (2017). The Ebola outbreak, 2013–2016: old lessons for new epidemics. Philos Trans R Soc Lond B Biol Sci 372: 20160297. doi: 10.1098/rstb.2016.0297.
5. World Health Organization (2014) Ebola outbreak 2014–2016 — West Africa. Available: https://www.who.int/emergencies/situations/ebola-outbreak-2014-2016-West-Africa. Accessed: 26 June 2026.
6. World Health Organization (2026). Ebola disease caused by Bundibugyo virus, Democratic Republic of the Congo & Uganda. Available: https://www.who.int/emergencies/disease-outbreak-news/item/2026-DON605. Accessed: 26 June 2026.
7. Pathoplexus (2026). Orthoebolavirus Sudan. Available: doi: 10.62599/PP_SS_2245.1. Accessed: 6 June 2026.
8. Pathoplexus (2026). Orthoebolavirus Zaire. Available: doi: 10.62599/PP_SS_2246.1. Accessed: 6 June 2026.
9. Pathoplexus (2026). Orthoebolavirus Bundibugyo. Available: doi: 10.62599/PP_SS_2212.1. Accessed: 6 June 2026.
10. Dutt M, Wayengera M, Martinez GS, Kumar A, Nsawotebba A, Ssewanya I, Nabadda S, Bakamutumaho B, Kirenga B, Olaro C, Muttamba W, Byakika P, Aceng JRO, Atwiine D, Bosa HK, Kelvin D, Sabiiti W (2026) Remdesivir triphosphate as a potential repurposed drug against the emerging Bundibugyo Ebolavirus 2026. J Infect 93: 106788. doi: 10.1016/j.jinf.2026.106788.
11. Ayitewala A, Amuri-Aziza A, Nsawotebba A, Wawina-Bokalanga T, Ssewanyana I, Nabadda S, Kinganda-Lusamaki E, Kanyerezi S, Zalwango VN, Mukadi-Bamuleka D, Mulopo-Mukanya N, Luakanda-Ndelemo G, Kumar A, Bosa HK, Sabiiti W, Muttamba W, Bakamutumaho B, Kirenga B, Dutt M, Martinez GS, Kelvin D, Olaro C, Aceng JRO, Atwiine D, Mbala-Kingebeni P, Wayengera M (2026) Bundibugyo ebolavirus of the 2026 Ebola outbreak in Uganda and DR Congo is a new variant. Lancet 407: S0140-6736(26)01079-2 (in press).
12. Nsawotebba A, Ssewanyana I, Ayitewala A, Wayengera M, Kanyerezi S, Kabazzi J, Nakintu V, Ssemaganda A, Kisuule R, Ssekyondwa S, Oundo HR, Kasujja R, Ayebare A, Balinandi S, Pimundu G, Muyigi T, Byekwaso GM, Mugerwa I, Kakooza F, Namboozo S, Namakula A, Kyokushaba J, Joloba M, Mutamba W, Bbuye M, Kirenga B, Kaleebu P, Sabiiti W, Bahatungire RR, Achola C, Ikoba S, Muruta A, Lubwama B, Kagirita A, Bosa HK, Aceng JR, Nabadda S, Olaro C, Atwine D (2026) Clinical profile and genomic characterization of the 2026 Bundibugyo virus index case in Uganda. Nat Med. doi: 10.1038/s41591-026-04510-7.
13. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30: 772–780. doi: 10.1093/molbev/mst010.
14. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R, Teeling E (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37: 1530–1534. doi: 10.1093/molbev/msaa015.
15. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR (2015) Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res 43: e15. doi: 10.1093/nar/gku1196.
16. Sagulenko P, Puller V, Neher RA (2018) TreeTime: maximum-likelihood phylodynamic analysis. Virus Evol 4: vex042. doi: 10.1093/ve/vex042.
17. Deen GF, Broutet N, Xu W, Knust B, Sesay FR, McDonald SLR, Ervin E, Marrinan JE, Gaillard P, Habib N, Liu H, Liu W, Thorson AE, Yamba F, Massaquoi TA, James F, Ariyarajah A, Ross C, Bernstein K, Coursier A, Klena J, Carino M, Wurie AH, Zhang Y, Dumbuya MS, Abad N, Idriss B, Wi T, Bennett SD, Davies T, Ebrahim FK, Meites E, Naidoo D, Smith SJ, Ongpin P, Malik T, Banerjee A, Erickson BR, Liu Y, Liu Y, Xu K, Brault A, Durski KN, Winter J, Sealy T, Nichol ST, Lamunu M, Bangura J, Landoulsi S, Jambai A, Morgan O, Wu G, Liang M, Su Q, Lan Y, Hao Y, Formenty P, Ströher U, Sahr F (2017). Ebola RNA persistence in semen of Ebola virus disease survivors — final report. N Engl J Med 377: 1428–1437. doi: 10.1056/NEJMoa1511410.
18. Vetter P, Fischer WA, Schibler M, Jacobs M, Bausch DG, Kaiser L (2016). Ebola virus shedding and transmission: review of current evidence. J Infect Dis 4: S177–S184. doi: 10.1093/infdis/jiw254.
19. van Vuren PJ, Ladner JT, Grobbelaar AA, Wiley MR, Lovett S, Allam M, Ismail A, le Roux C, Weyer J, Moolla N, Storm N, Kgaladi J, Sanchez-Lockhart M, Conteh O, Palacios G, Paweska JT (2019) Phylodynamic analysis of Ebola virus disease transmission in Sierra Leone. Viruses 11: 71. doi: 10.3390/v11010071.
20. Goldstein T, Anthony SJ, Gbakima A, Bird BH, Bangura J, Tremeau-Bravard A, Belaganahalli MN, Wells HL, Dhanota JK, Liang E, Grodus M, Jangra RK, DeJesus VA, Lasso G, Smith BR, Jambai A, Kamara BO, Kamara S, Bangura W, Monagin C, Shapira S, Johnson CK, Saylors K, Rubin EM, Chandran K, Lipkin WI, Mazet JAK (2018) The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses. Nat Microbiol 3: 1084–1089. doi: 10.1038/s41564-018-0227-2.