nature portfolio

Peer Review File

Emergence of a new Bundibugyo virus variant in the 2026 outbreak in the Democratic Republic of the Congo and Uganda

Corresponding Author: Dr Tony Wawina-Bokalanga

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This report presents a phylogenetic analysis of n=21 BDBV genomes in the context of BDBV genomes from the 2012 and 2007 outbreaks. This analysis concludes that this outbreak very likely originated from a novel zoonotic introduction/spillover, rather than reemergence of virus from an earlier outbreak. This conclusion is well-supported by the lower left phylogeny in Figure 1. Regarding the phylogenetic relatedness of BDBV genomes from the current outbreak, there are not many take-aways noted, although a paragraph is focused on not misinterpreting the apparent pattern of Ugandan sequences appearing in 3 distinct viral clades, even though epidemiological evidence indicates that cases in Uganda derived from a single introduction from the DRC.

Major comments:

Figure 1/text lines 144-147: I agree with the interpretation of the analyses shown in the lower left panel. However, it would be helpful to know how many genomes from the 2012 and the 2007 outbreaks are used in Figure 1 lower left panel. Please provide this information and the accession numbers of these genomes for reproducibility in a separate Table. Are these genomes high-quality as well, according to the criteria on lines 142-143?

I am confused about the entries in Table 1. Do they include the clinical specimens from the n=19 suspected and probable cases of EVD (line 109) or do they contain the 22 BDBV genomes (line 130)? Do these samples overlap or are they mutually exclusive? The 21 different patients alluded to on line 142 I assume are all but one from line 130 (the DRC sample that was lower quality)? I thought the Table would have n=22 rows, with one of the genomes not submitted. Instead, there are n=27? Are some sequences from the same individual, using a different sampling approach (e.g. blood vs oropharyngeal swab?) Could the Table also list sampling approach/tissue sampled? Also, why waiting for submission for a handful of them?

Figure 1 right panel phylogeny/text lines 149-157: I am concerned about the three subclusters that are interpreted as unresolved polytomies. If this is the case, it is not visually clearly apparent. I understand that limited genetic divergence can be an issue, but there are other ways to visualize relationships between genomes without a phylogeny in that case—why not show (and interpret) a haplotype network instead (e.g., Murcia et al. http://jvi.asm.org/cgi/doi/10.1128/JVI.00112-10 Fig 2B)? Does such a network provide some evidence to support (ore refute) a single introduction into Uganda? Another benefit of showing a haplotype network would be that all genetic variation can be shown on the network (about 20 mutations in all?) Currently, I am unclear as to what we learn from the phylogeny on the right. (There is also no mention of any insight/takeaway in the abstract for this right panel result.) Finally, it is difficult to interpret which samples are from DRC and which from Uganda. Perhaps make the marker different, e.g., circles (of different colors) for DRC and diamonds (of different colors) for Uganda? (Although, again, a phylogeny may not be the best visual to use here.)

Given limited amount of genetic variation, it would be helpful if it were characterized, at least to some extent. How many polymorphic sites exhibit synonymous variation? How many nonsynonymous variation?

Figure 1 – title of figure does not reflect all panels, just the right panel. Also, last line of figure legend refers only to lower left

panel I think. Legend text should therefore be modified for readability and accuracy.

Minor comments:

The abstract and introduction introduce a lot of acronyms, several of which are only used 2-3 times in the remainder of the text. This includes ED (Ebola Disease) and EVD (Ebola Virus Disease). I suggest for improved readability to remove these acronyms. BVD and BDBV are used much more often so they should of course be kept.

Line 83: 'Besides strengthening surveillance systems and community engagement, and establishing Ebola treatment centers, our report …' Reword/rephrase for accuracy? The report itself does not strengthen surveillance systems or establish Ebola treatment centers.

Line 128: 'PCR-confirmed ,including' -> 'PCR-confirmed cases, including'

Line 142: clarify that these n = 21 genomes are those from the n=22 samples on line 130 (minus 1 DRC sample), if that is the case

Refer to Figure 1 on line 147 (after 'outbreaks') rather than a sentence later?

Reviewer #2

(Remarks to the Author)
This manuscript provides a timely and important genomic characterization of the ongoing 2026 Bundibugyo virus disease (BVD) outbreak in the Democratic Republic of the Congo and Uganda. A major strength of the study is the rapid generation and sharing of genomic data during an active outbreak, including 22 newly generated Bundibugyo virus genomes from two countries. The work demonstrates strong regional collaboration and highlights the value of genomic surveillance for outbreak response. The phylogenetic analyses show that viruses associated with the 2026 outbreak form a distinct, well-supported lineage separate from those responsible for the 2007 Uganda and 2012 DRC outbreaks, providing important new insights into the evolutionary history of Bundibugyo virus. The authors also appropriately acknowledge the limited phylogenetic resolution among Ugandan genomes and avoid overinterpreting unresolved relationships as evidence of multiple introductions. Furthermore, the diagnostic observations are highly relevant, demonstrating that genomic sequencing successfully confirmed infections that were missed or returned invalid results using some existing molecular assays.
Despite these strengths, several limitations reduce the impact of the study. The principal conclusion that the outbreak represents a new zoonotic spillover event is stronger than the available evidence supports. While the phylogenetic data clearly indicate that the 2026 viruses are distinct from previously sequenced outbreaks, alternative explanations such as circulation in unsampled reservoirs, undetected transmission chains, or emergence from an unsampled lineage cannot be excluded. The data therefore support the interpretation that the outbreak is consistent with, rather than definitive evidence of, a novel spillover event. In addition, the study relies exclusively on maximum-likelihood phylogenetic reconstruction and lacks molecular clock analyses that could estimate divergence times and provide stronger support for the evolutionary origins of the lineage. The phylogenetic methods are relatively limited for a study centered on evolutionary inference, with no assessment of temporal signal, phylogeographic reconstruction, or evolutionary rate estimation.
The manuscript would also benefit from greater integration of epidemiological and genomic data. Although epidemiological investigations are mentioned, the study does not explore transmission networks, geographic spread, or movement patterns in detail. Likewise, the genomic characterization remains largely descriptive, with little discussion of lineage-defining mutations, amino acid changes, or potential implications for diagnostics, therapeutics, or vaccine development. The representativeness of the sequenced samples is also unclear given the large number of reported cases relative to the number of genomes analyzed. Finally, there are several minor issues, including apparent citation errors related to references supporting the declaration of public health emergencies and some inconsistencies in terminology throughout the manuscript.
Overall, this is a valuable and highly relevant outbreak report that makes an important contribution to the understanding of Bundibugyo virus emergence and outbreak response. The rapid generation and sharing of genomic data during an ongoing outbreak is commendable and provides an important resource for public health authorities and the scientific community. However, I urge the authors to complement the current work with more in-depth evolutionary analyses. In particular, molecular clock dating, assessment of temporal signal, phylogeographic reconstruction, and a detailed characterization of lineage-defining mutations including selection analyses would substantially strengthen the conclusions regarding the origins, emergence, and transmission dynamics of the 2026 outbreak lineage. Such analyses would provide a more robust framework for evaluating whether the observed genetic diversity is consistent with a recent zoonotic spillover event or alternative epidemiological scenarios. With revisions that temper conclusions regarding zoonotic spillover, expand the evolutionary analyses, and better integrate epidemiological context, the manuscript would provide a stronger and more comprehensive account of this significant public health event.

Reviewer #3

(Remarks to the Author)
Amuri-Aziza et al. conducted a molecular surveillance study during the 2026 Bundibugyo Virus Disease (BVD) outbreak. They generated 22 BVD virus genomes from patient samples collected during the outbreak and found these sequences to be distinct from the virus strains responsible for the 2007 and 2012 outbreaks. This key finding suggests a novel zoonotic emergence rather than a resurgence of prior strains. The sequencing and analytical methods employed in this study were well-executed, with strong writing and organization throughout their presentation. The interpretation and conclusion are generally robust (except the three subclusters issue I mentioned below). This timely and important research could provide

valuable insights for controlling the ongoing BVD outbreak.

Below are several suggestions for consideration:
1. Line 149:

The sentence: “While the eight BDBV genomes from Uganda seem to generate three subclusters, …”
Upon examining the tree lineage generated by the 2026 outbreak sequences (Figure 1, right panel), only two subclusters are apparent from the Uganda sequences. There were only seven sequences I could see were from 2026 outbreak in Uganda. Could this be clarified further? Is it possible that the three subclusters include a subcluster from the 2007 outbreak that is represented by the orange triangle in the left panel? If so, it means some 2026 outbreak infections were caused by resurgence of prior outbreak strains that have been circulating?

Although the 2007 and 2012 outbreaks are not the focus of this study, the simplified presentation of tree lineages (e.g., using triangle representations for lineages) can sometimes make interpretation difficult. For instance, the tree branch structure and exact number of sequences forming these lineages are unclear. Would it be possible to provide a full display of the phylogenetic tree, either in the supplementary section or even in the main text? This would enhance clarity and interpretation.

2. Figure 1:

Does the phylogenetic tree include any measures of topological support (e.g., bootstrap values)? Displaying these would enhance confidence in the inferred subclusters and provide crucial information about their robustness.

Can you provide details about the specific mutations that occurred in the virus genome, both during the 2026 outbreak and those that define the emergence of the 2026 outbreak lineage?

Is it feasible to estimate the timescale of the 2026 BVD outbreak? How strong is the temporal signal in the sequence dataset to allow for reliable molecular clock analysis?

3. Map Representation (Figure 1, Top Left Panel):

The variable size of location names on the map, such as Kampala appearing much larger than Wakiso, seems inconsistent. Standardizing the size of location names would improve readability and presentation quality.

Additionally, it would be useful to use consistent coloring for the map dots, aligned with the map's color legend on the right for clarity and interpretability.

[redacted]

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I am satisfied with the edits made. A couple of minor additional comments:
- The introduction is still very acronym-heavy.
- GenBank accession NC_014373 being used as the reference genome on which to call mutations. Would be helpful (and important) to add a sentence as to why this one was chosen as the reference genome.
- In the Table providing the newly sequenced genomes, under the column GC content, sometimes the percentage is specified as XX.X% and sometimes XX,X%. Standardize.

Reviewer #2

(Remarks to the Author)
The authors provided a comprehensive response to my queries.

Reviewer #3

(Remarks to the Author)
The authors have addressed my previous comments, and the revised manuscript appears to be substantially improved . I have only a few minor suggestions for further refinement .

1. While Figure 1 has been considerably improved, particularly in terms of colour consistency, I would further suggest ensuring that the symbol shapes in the right-hand tree are fully consistent with those used on the map . At present, taxa in the right-hand tree are shown as circles (DRC) and diamond (Uganda); however, the labels of samples from Uganda on the map notation should also be represented as diamonds to match the tree taxa shape.

2. Although the authors justify the collapsed BDBV tree as a clearer way to highlight relationships among outbreaks and to support the proposed novel zoonotic spillover event, I still recommend including the complete, uncollapsed tree in the supplementary materials for readers who are interested in the full phylogenetic context . In that supplementary full tree, the authors could consider using different branch-length scales for within-outbreak and between-outbreak clades to improve readability without losing either resolution or overall phylogenetic structure.

3. I also suggest adding a “#” symbol before the node labels to distinguish them more clearly from bootstrap values and thereby avoid potential confusion.

Reviewer #1:

This report presents a phylogenetic analysis of n=21 BDBV genomes in the context of BDBV genomes from the 2012 and 2007 outbreaks. This analysis concludes that this outbreak very likely originated from a novel zoonotic introduction/spillover, rather than reemergence of virus from an earlier outbreak. This conclusion is well-supported by the lower left phylogeny in Figure 1. Regarding the phylogenetic relatedness of BDBV genomes from the current outbreak, there are not many take-aways noted, although a paragraph is focused on not misinterpreting the apparent pattern of Ugandan sequences appearing in 3 distinct viral clades, even though epidemiological evidence indicates that cases in Uganda derived from a single introduction from the DRC.

Response: We thank the reviewer for their thorough evaluation and critical assessment, which helped us further improve our work. We have now expanded the analysis within the right panel high-resolution phylogenetic presentation by (i) reconstructing the detailed evolutionary history of all 21 BDBV genomes (ii) incorporating an ADAR-resolved phylogeny indicating ADAR mutations as was seen in past human Ebola virus disease outbreaks, and (iii) describing the individual mutations contributing to this evolution (Extended Data Table 3. Mutational profile of the 2026 BDBV genomes with nucleotide coordinates relative to BDBV reference genome). For your reference, we have included the revised figure below:

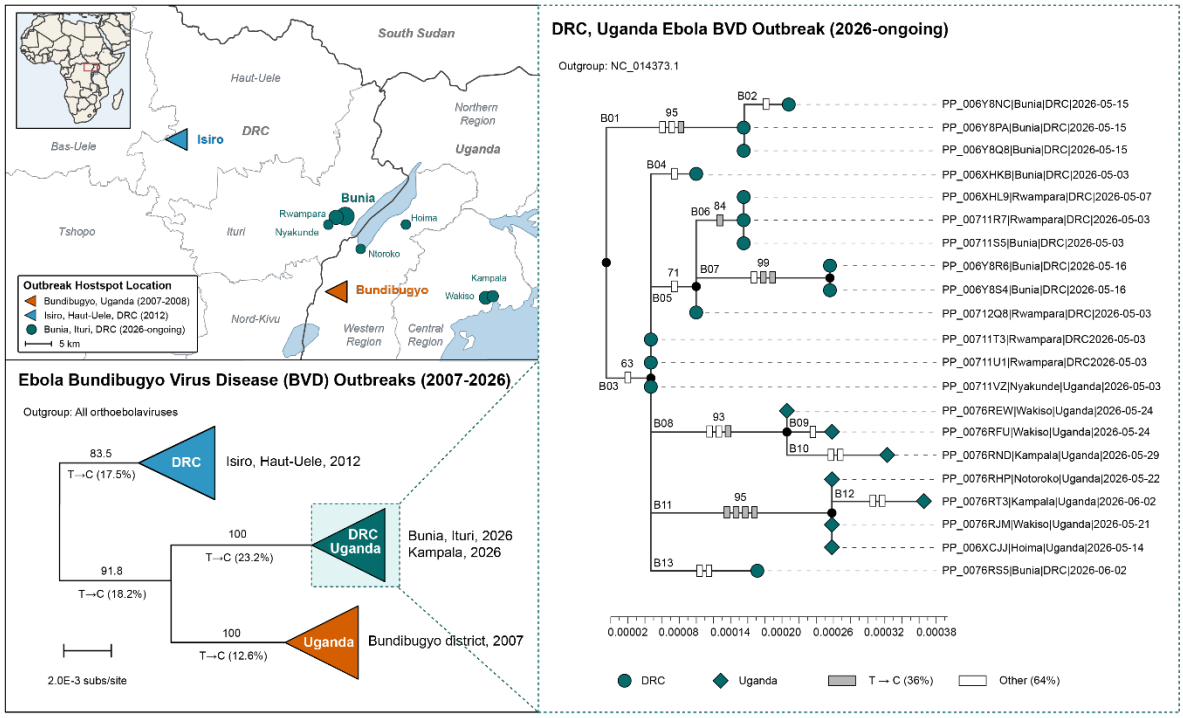


Figure 1 | Geographic distribution and phylogenetic analyses of Bundibugyo virus (BDBV) genomes. **Upper left panel:** Map of eastern DRC and Uganda showing all known BVD outbreak hotspot locations. **Lower left panel:** Maximum-likelihood phylogeny of all known BVD outbreaks inferred using the +gamma model in IQTree2. Ultrafast bootstrap support values are shown for each branch. The scale is given in units of substitutions per site. The current outbreak is highlighted in dark green, previous outbreaks in orange (Uganda, 2007) and blue (DRC, 2012). **Right panel:** Maximum-likelihood phylogeny of 21 high-quality 2026 BDBV genomes using the same parameters as described above. Mutations identified on each branch are indicated (ADAR-like and other), and branch IDs (i.e., 1-13) correspond to those listed in Extended Data Table 3 (Mutational analysis), providing a high-

resolution view of the evolutionary relationships and mutations acquired during the 2026 outbreak. PP_006Y8R6 and PP_006Y8S4 represent two samples collected from the same patient.

We have also revised the text as follows:

***Lines 157-166*:** Within the 2026 BDBV outbreak, phylogenetic analysis confirms cross-border transmission between DRC and Uganda (Figure 1, right panel). Importantly, available epidemiological investigations indicate that all included Ugandan cases belong to a single transmission chain, highlighting that the apparent separation of sequences in the phylogeny should not be overinterpreted as evidence for independent introductions. Additional genome data may help resolve these relationships in the future. Mutational analysis identified 25 mutations in total within the 2026 BDBV genome cluster, of which 9/25 (36%) showed patterns consistent with Adenosine Deaminase Acting on RNA (ADAR)-mediated editing. ADAR mutations are a signature of ebolavirus transmission in the human population. Of these 25 mutations, five resulted in non-synonymous substitutions (further detailed in Extended Data Table 3).

Major comments:

1. Figure 1/text lines 144-147: I agree with the interpretation of the analyses shown in the lower left panel. However, it would be helpful to know how many genomes from the 2012 and the 2007 outbreaks are used in Figure 1 lower left panel. Please provide this information and the accession numbers of these genomes for reproducibility in a separate Table. Are these genomes high-quality as well, according to the criteria on lines 142-143?

Response 1: All publicly available near-complete Bundibugyo virus genomes (defined as having > 80% horizontal genomes coverage) were included for constructing the maximum-likelihood phylogeny shown in Figure 1 (n=34). This dataset includes genomes from both the 2007 (n=11) and 2012 (n=23) outbreaks. We have included the Extended Data Table 2, which lists previous BDBV genomes used in this study. Extended Data Table 2 includes the NCBI GenBank accession numbers (Accession), date of sample collection (Date), country of origin (Country), associated BVD outbreak (Outbreak), first author(s) and submitting institution (Institution). An example of the Table is provided below:

Accession	Date	Country	Outbreak	Authors	Institution
FJ217161.1	Nov. 2007	Uganda	BDBV-2007	Towner, J. et al	CDC

2. I am confused about the entries in Table 1. Do they include the clinical specimens from the n=19 suspected and probable cases of EVD (line 109) or do they contain the 22 BDBV genomes (line 130)? Do these samples overlap or are they mutually exclusive? The 21 different patients alluded to on line 142 I assume are all but one from line 130 (the DRC 6

Response 2: Table 1 includes only data from PCR-confirmed cases of Bundibugyo virus disease. To avoid any confusion, we clarified this point in the title as follows:

Table 1. Socio-demographic, clinical symptoms, molecular testing results of PCR-confirmed BVD patients (n=22), as well as BDBV genome coverage.

3. I thought the Table would have n=22 rows, with one of the genomes not submitted. Instead, there are n=27? Are some sequences from the same individual, using a different sampling approach (e.g. blood vs oropharyngeal swab?).

Response 3: Corrected in the updated Table 1.

4. Could the Table also list sampling approach/tissue sampled? Also, why waiting for submission for a handful of them?

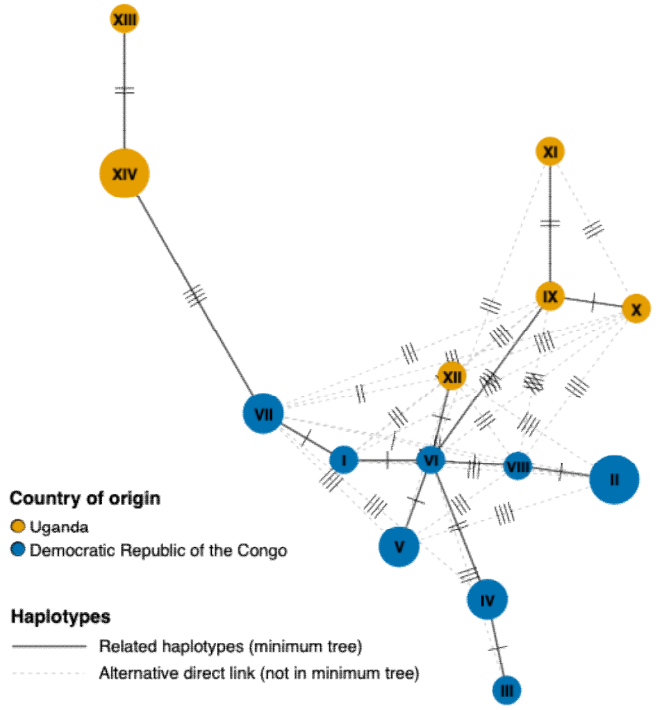
Response 4: All eight genomes that were previously pending submission have now been submitted, and their corresponding Pathoplexus IDs have been added to the table. Regarding the sampling approach, we have included the following paragraph in the revised manuscript:

***Lines 221-225*:** In this study, blood samples were collected by venipuncture into EDTA tubes, except for one case from Uganda, for which posthumous EDTA whole blood was available. Oral swab specimens were collected among deceased individuals by gently rubbing sterile dry swabs against the buccal mucosa (the inner surfaces of the cheeks) to collect oral fluids and cellular material.

5. Figure 1 right panel phylogeny/text lines 149-157: I am concerned about the three subclusters that are interpreted as unresolved polytomies. If this is the case, it is not visually clearly apparent. I understand that limited genetic divergence can be an issue, but there are other ways to visualize relationships between genomes without a phylogeny in that case - why not show haplotype network (e.g., Murcia et al. http://jvi.asm.org/cgi/doi/10.1128/JVI.00112-10 Fig 2B)? Does such a network provide some evidence to support (ore refute) a single introduction into Uganda? Another benefit of showing a haplotype network would be that all genetic variation can be shown on the network (about 20 mutations in all?).

Response 5: We thank the reviewer for this suggestion regarding alternative visualizations. We evaluated whether a haplotype network would improve the resolution of the observed relationship. Specifically, we generated a fully connected haplotype network using the same Bundibugyo virus aligned sequences (n=21), which were used to construct the phylogenetic tree (See Figure below). The network was generated using the *Pegas* and *ape* packages in R, with each node representing a unique viral genome sequence (haplotype) and connections representing the number of mutations separating haplotypes. Solid lines represent the minimum spanning network, while dashed lines indicate alternative possible connections between haplotypes that are not part of the minimum spanning solution but remain genetically plausible. However, due to limited genetic divergence among the sequences, the haplotype network showed similarly unresolved relationships and **did not provide additional clarity beyond the phylogenetic tree presented in Figure 1**.

As the fully connected haplotype network does not provide sufficient additional resolution to support or reject a single introduction event into Uganda, we therefore opted not to include it in the manuscript. Our interpretation of these results is consistent with the comments of the second reviewer, who noted that we appropriately acknowledge the limited phylogenetic resolution among Ugandan genomes and avoid overinterpreting unresolved relationships as evidence for multiple introductions.



Fully connected haplotype network of Bundibugyo virus sequences showing all possible connection between haplotypes.

6. Currently, I am unclear as to what we learn from the phylogeny on the right. (There is also no mention of any insight/takeaway in the abstract for this right panel result).

Response 6: We kindly refer the reviewer to our first response. We have now revised the abstract as follows:

***Lines 85-87*:** These genomes form a well-supported phylogenetic cluster separate from BDBV variants associated with the 2007 and 2012 outbreaks, together with evidence for ADAR editing in the viral genomes.

7. Finally, it is difficult to interpret which samples are from DRC and which from Uganda. Perhaps make the marker different, e.g., circles (of different colors) for DRC and diamonds (of different colors) for Uganda? (Although, again, a phylogeny may not be the best visual to use here).

Response 7: We agree with the reviewer and have implemented the suggested changes to improve clarity of figure 1, applying circles and diamonds (please see the revised figure 1).

8. Given limited amount of genetic variation, it would be helpful if it were characterized, at least to some extent. How many polymorphic sites exhibit synonymous variation? How many nonsynonymous variations?

Response 8: We agree with the reviewer that additional characterization of the observed genetic variation would strengthen the manuscript. We have therefore added Extended Data Table 3, which summarizes the mutational profile of the 2026 BDBV genomes. This analysis identified 25 mutations, including five non-synonymous amino acid substitutions (*lines 162-166*).

9. Figure 1 – title of figure does not reflect all panels, just the right panel. Also, last line of figure legend refers only to lower left panel, I think. Legend text should therefore be modified for readability and accuracy.

Response 9: We agree with the reviewer and have revised the legend of Figure 1 as follows:

*"**Figure 1** | Geographic distribution and phylogenetic analyses of Bundibugyo virus (BDBV) genomes. **Upper left panel:** Map of eastern DRC and Uganda showing all known BVD outbreak hotspot locations. **Lower left panel:** Maximum-likelihood phylogeny of all known BDBV outbreaks inferred using the +gamma model in IQTree2. Ultrafast bootstrap support values are shown for each branch. The scale is given in units of substitutions per site. The current outbreak is highlighted in green, previous outbreaks in orange (Uganda, 2007) and blue (DRC, 2012). **Right panel:** Maximum-likelihood phylogeny of 21 high-quality 2026 BDBV genomes using the same parameters as described above. Mutations identified on each branch are indicated (ADAR-like and other), and branch IDs (i.e., 1- 13) correspond to those listed in Extended Data 3 (Mutational analyses), providing a high-resolution view of the evolutionary relationships and mutations acquired during the 2026 outbreak. PP_006Y8R6 and PP_006Y8S4 represent two samples collected from the same patient."*

Minor comments:

10. The abstract and introduction introduce a lot of acronyms, several of which are only used 2-3 times in the remainder of the text. This includes ED (Ebola Disease) and EVD (Ebola Virus Disease). I suggest for improved readability to remove these acronyms. BVD and BDBV are used much more often so they should of course be kept.

Response 10: We have removed unnecessary acronyms while retaining BVD (Bundibugyo Virus Disease) and BDBV (Bundibugyo virus), which are extensively used through the manuscript.

11. Line 83: 'Besides strengthening surveillance systems and community engagement, and establishing Ebola treatment centers, our report …' Reword/rephrase for accuracy? The report itself does not strengthen surveillance systems or establish Ebola treatment centers.

Response 11: We have revised the sentence to clarify that these activities represent ongoing public health efforts rather than actions conducted during the outbreak.

***Lines 89-93*:** Besides ongoing efforts in strengthening surveillance systems, community engagement, establishing Ebola treatment centers, and developing targeted medical countermeasures; our report advocates to specifically increase decentralized laboratory diagnostics capacity, with BDBV-compatible assays, including genomic sequencing capacity, for limiting further outbreak expansion, timely detection and control of future outbreaks.

12. Line 128: 'PCR-confirmed ,including' -> 'PCR-confirmed cases, including'.

Response 12: Thanks. "Cases" has been added in the revised manuscript (*line 136*).

13. Line 142: clarify that these n = 21 genomes are those from the n=22 samples on line 130 (minus 1 DRC sample), if that is the case.

Response 13: After careful revision, we noticed an error in the genome coverage value of sample 26FHV058 with Pathoplexus ID "PP_006Y8ME" (Table 1 and Extended Data Table 1). The reported coverage value should have been 25%; we have corrected this in the table.

***Lines 149-151*:** High-quality BDBV genomes from 21 patients (14 from DRC and seven from Uganda), with a minimum mean depth coverage of 45x and genome coverage >80% (Extended Data Table 1), were used to infer the phylogenetic tree.

14. Refer to Figure 1 on line 147 (after 'outbreaks') rather than a sentence later?

Response 14: Revised as suggested. We have moved the reference to Figure 1 to the relevant sentence as follows:

***Lines 151-156*:** Phylogenetic analysis, including contextual BDBV full genomes from previous outbreaks (Extended Data Table 2), demonstrated that the 2026 BDBV genomes form a distinct, well-supported cluster separate from viruses associated with the 2007 and 2012 outbreaks (Figure 1, left lower panel). Given the distinct phylogenetic placement these findings are consistent with a zoonotic spillover event rather than a resurgence from previously known BDBV variants.

Reviewer #2:

This manuscript provides a timely and important genomic characterization of the ongoing 2026 Bundibugyo virus disease (BVD) outbreak in the Democratic Republic of the Congo and Uganda. A major strength of the study is the rapid generation and sharing of genomic data during an active outbreak, including 22 newly generated Bundibugyo virus genomes from two countries. The work demonstrates strong regional collaboration and highlights the value of genomic surveillance for outbreak response. The phylogenetic analyses show that viruses associated with the 2026 outbreak form a distinct, well-supported lineage separate from those responsible for the 2007 Uganda and 2012 DRC outbreaks, providing important new insights into the evolutionary history of Bundibugyo virus. The authors also appropriately acknowledge the limited phylogenetic resolution among Ugandan genomes and avoid overinterpreting unresolved relationships as evidence of multiple introductions. Furthermore, the diagnostic observations are highly relevant, demonstrating that genomic sequencing successfully confirmed infections that were missed or returned invalid results using some existing molecular assays.

Response : We sincerely thank the reviewer for their positive evaluation of our manuscript and for recognizing the importance of the rapid genomic characterization, regional collaboration, and diagnostic findings presented in this study. We have carefully considered each of the comments and suggestions provided and have revised the manuscript accordingly. Below, we provide a detailed point-by-point response describing how each remark has been addressed.

1. Despite these strengths, several limitations reduce the impact of the study. The principal conclusion that the outbreak represents a new zoonotic spillover event is stronger than the available evidence supports. While the phylogenetic data clearly indicate that the 2026 viruses are distinct from previously sequenced outbreaks, alternative explanations such as circulation in unsampled reservoirs, undetected transmission chains, or emergence from an unsampled lineage cannot be excluded. The data therefore support the interpretation that the outbreak is consistent with, rather than definitive evidence of, a novel spillover event.

Response 1: We agree with the reviewer that our original wording was too definitive regarding the origin of the outbreak. We have therefore revised the text to more accurately reflect the findings.

***Lines 151-156*:** Phylogenetic analysis, including contextual BDBV full genomes from previous outbreaks (Extended Data Table 2), demonstrated that the 2026 BDBV genomes form a distinct, well-supported cluster separate from viruses associated with the 2007 and 2012 outbreaks (Figure 1, left lower panel). Given the distinct phylogenetic placement, these findings are consistent with a zoonotic spillover event rather than a resurgence from previously known BDBV variants.

2. In addition, the study relies exclusively on maximum-likelihood phylogenetic reconstruction and lacks molecular clock analyses that could estimate divergence times and provide stronger support for the evolutionary origins of the lineage. The phylogenetic methods are relatively limited for a study centered on evolutionary inference, with no assessment of temporal signal, phylogeographic reconstruction, or evolutionary rate estimation.

Response 2: The main goal of this study was to rapidly inform the public health stakeholders and the scientific community about the BDBV species involved and its likely origin (i.e., a new zoonotic spillover event versus a resurgence from previous outbreaks in the region), as well as to reconstruct the early diagnostic events leading to species confirmation, highlighting the relevance of viral whole genome sequencing. However, we have now explained this more and advocated indeed for the relevance of generating more genomic data.

***Lines 175-184*:** While centralized laboratory molecular diagnostic testing and genomic sequencing enabled case confirmation and virus characterization, increasing similar de-centralized capacity might help to detect outbreaks at even earlier stages. Thus, rapid deployment of sustainable decentralized molecular diagnostic testing and genomic sequencing capacity should further decrease turn-around times for pathogen detection, patient management, outbreaks containment, and better prepare for future outbreaks. Decentralized sequencing capacity will enable the generation of additional genomes needed to accurately estimate the date of the most recent common ancestor of the 2026 BVD outbreak. In addition, near-real-time genomic surveillance will assist in early identification of viral spread dynamics and viral evolution affecting outbreak measures.

3. The manuscript would also benefit from greater integration of epidemiological and genomic data. Although epidemiological investigations are mentioned, the study does not explore transmission networks, geographic spread, or movement patterns in detail.

Response 3: We thank the reviewer for raising this point. The lack of detailed epidemiological data linked to individual cases limited our ability to infer transmission chains or analyze transmission networks, geographic spread and movement patterns. As noted in Response #2, expanding decentralized laboratory capacity will improve geographic coverage and enable more detailed integration of genomic and epidemiological data in future studies.

4. Likewise, the genomic characterization remains largely descriptive, with little discussion of lineage-defining mutations, amino acid changes, or potential implications for diagnostics, therapeutics, or vaccine development.

Response 4: One of the objectives of this early outbreak report is to make the genomic data available for the broader research community as a resource for further analyses, including those suggested by the reviewer. To avoid delaying release of this important resource, we considered such analyses beyond the scope of the current study. We therefore refer the reviewer to our final conclusion statement (lines 199-202), which was intended to stimulate such follow-up research:

Lines 190-193: Timely release of sequences of BDBV variant causing the 2026 BVD epidemic in DRC and Uganda is relevant to public health, providing an important resource for the design or adaptation of specific medical countermeasures (vaccines and therapeutics) as well as molecular diagnostic tests.

Regarding lineage-defining mutations, we have provided a more detailed description of the phylogenetic findings and incorporated an analysis of observed mutations. The revised text is as follows:

Lines 162-166: Mutational analysis identified 25 mutations in total within the 2026 BDBV genome cluster, of which 9/25 (36%) showed patterns consistent with Adenosine Deaminase Acting on RNA (ADAR)-mediated editing. ADAR mutations are a signature of ebolavirus transmission in the human population. Of these 25 mutations, five resulted in non-synonymous substitutions (further detailed in Extended Data Table 3).

5. The representativeness of the sequenced samples is also unclear given the large number of reported cases relative to the number of genomes analyzed.

Response 5: We acknowledge the reviewer's concern that the sequenced samples may not fully represent the geographic distribution of currently reported cases while the outbreak is continuously expanding. However, the primary aim of this study was to rapidly report these genomic data and reconstruct the early diagnostic events to determine whether the outbreak resulted from a new zoonotic spillover event, or a resurgence linked to previous outbreaks. For this objective, we focused on the very initial samples collected through outbreak epidemiological investigations that led to outbreak declarations in both affected countries specifically.

6. Finally, there are several minor issues, including apparent citation errors related to references supporting the declaration of public health emergencies and some inconsistencies in terminology throughout the manuscript.

Response 6: Thank you for bringing these minor issues to our attention. The correct references have now been included in the revised manuscript.

Lines 131-134: A Public Health Emergency of International Concern (PHEIC) (7.)and a Public Health Emergency of Continental Security (PHECS) (8.) were declared by the World Health Organization and Africa Centres for Disease Control and Prevention (Africa CDC) on 17 May 2026 and 18 May 2026, respectively.

7. Overall, this is a valuable and highly relevant outbreak report that makes an important contribution to the understanding of Bundibugyo virus emergence and outbreak response. The rapid generation and sharing of genomic data during an ongoing outbreak is commendable and provides an important resource for public health authorities and the scientific community.

Response 7: We again appreciate the reviewer's positive feedback and recognition of the importance of rapid genomic data generation and sharing during outbreaks.

8. However, I urge the authors to complement the current work with more in-depth evolutionary analyses. In particular, molecular clock dating, assessment of temporal signal, phylogeographic reconstruction, and a detailed characterization of lineage-defining mutations including selection analyses would substantially strengthen the conclusions regarding the origins, emergence, and transmission dynamics of the 2026 outbreak lineage. Such analyses would provide a more robust framework for evaluating whether the observed genetic diversity is consistent with a recent zoonotic spillover event or alternative epidemiological scenarios. With revisions that temper conclusions regarding zoonotic spillover, expand the evolutionary analyses, and better integrate epidemiological context, the manuscript would provide a stronger and more comprehensive account of this significant public health event.

Response 8: We kindly refer the reviewer to our previous responses to comments 1,2 and 4. We believe that the revisions made in response to these comments have appropriately clarified the scope of the study and strengthened the evolutionary interpretations of the genomic findings.

Reviewer #3:

Amuri-Aziza et al. conducted a molecular surveillance study during the 2026 Bundibugyo Virus Disease (BVD) outbreak. They generated 22 BVD virus genomes from patient samples collected during the outbreak and found these sequences to be distinct from the virus strains responsible for the 2007 and 2012 outbreaks. This key finding suggests a novel zoonotic emergence rather than a resurgence of prior strains. The sequencing and analytical methods employed in this study were well-executed, with strong writing and organization throughout their presentation. The interpretation and conclusion are generally robust (except the three subclusters issue I mentioned below). This timely and important research could provide valuable insights for controlling the ongoing BVD outbreak.

Response: We sincerely thank the reviewer for their positive and encouraging comments. We appreciate their recognition of the quality of sequencing, analyses, and presentation, as well as the significance of our findings.

Below are several suggestions for consideration:

1. Line 149: The sentence: "While the eight BDBV genomes from Uganda seem to generate three subclusters, …". Upon examining the tree lineage generated by the 2026 outbreak sequences (Figure 1, right panel), only two subclusters are apparent from the Uganda sequences. There were only seven sequences I could see were from 2026 outbreak in Uganda. Could this be clarified further? Is it possible that the three subclusters include a subcluster from the 2007 outbreak that is represented by the orange triangle in the left panel? If so, it means some 2026 outbreak infections were caused by resurgence of prior outbreak strains that have been circulating?

Response 1: This is a keen observation. We re-evaluated the data and confirmed that there are two, rather than three, Ugandan subclusters. One genome (PP_0076RS5) was previously assigned to Uganda with an unknown location, which contributed to this confusion. Upon

further review, we found that the individual's place of residence was Bunia, DRC, although the sample was sequenced in Kampala, Uganda. This has now been corrected accordingly.

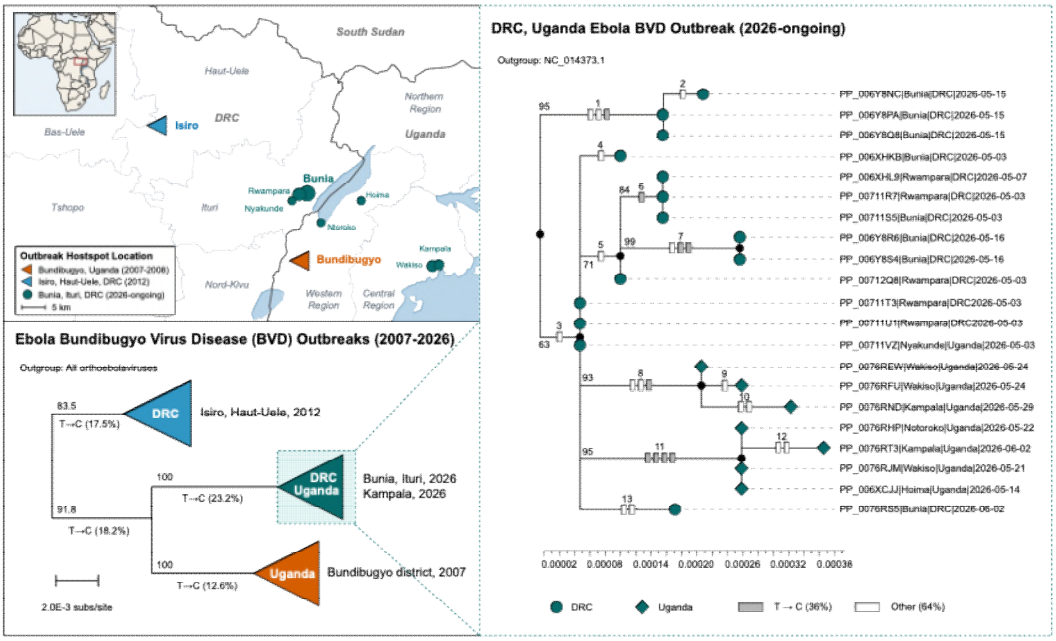
2. Although the 2007 and 2012 outbreaks are not the focus of this study, the simplified presentation of tree lineages (e.g., using triangle representations for lineages) can sometimes make interpretation difficult. For instance, the tree branch structure and exact number of sequences forming these lineages are unclear. Would it be possible to provide a full display of the phylogenetic tree, either in the supplementary section or even in the main text? This would enhance clarity and interpretation.

Response 2: We appreciate the reviewer's comment. The collapsed BDBV tree (Figure 1, lower left panel) was designed to highlight the evolutionary relationships between Bundibugyo virus disease outbreaks and provide support for the novel zoonotic spillover event. However, this level of resolution does not capture the fine-scale diversification within the 2026 outbreak, which appears as nearly vertical branching patterns in an uncollapsed tree. Therefore, we retain this panel for between-outbreak comparisons and use the right panel for higher-resolution within-outbreak descriptions. Furthermore, we have included Extended Data table 2 "Previously published BDBV genome sequences used in this study with NCBI GenBank accession (Accession), date of sample collection (Date), country of origin (Country), associated BVD outbreak (Outbreak), Authors, and submitting institution (Institution)", which includes relevant metadata for all BDBV sequences used in this study.

***Lines 151-154*:** Phylogenetic analysis, including contextual BDBV full genomes from previous outbreaks (Extended Data Table 2), demonstrated that the 2026 BDBV genomes form a distinct, well-supported cluster separate from viruses associated with the 2007 and 2012 outbreaks (Figure 1, left lower panel).

3. Figure 1: Does the phylogenetic tree include any measures of topological support (e.g., bootstrap values)? Displaying these would enhance confidence in the inferred subclusters and provide crucial information about their robustness.

Response 3: In line with this comment, we have now added bootstrap support values to the phylogenetic tree to allow for robustness assessment of the inferred (sub)clusters (See revised Figure 1 down below).



4. Can you provide details about the specific mutations that occurred in the virus genome, both during the 2026 outbreak and those that define the emergence of the 2026 outbreak lineage?

Response 4: Thank you for this comment. We further refined Figure 1 by adding the acquired mutations (ADAR-like and others) to each branch of the phylogenetic tree. Each branch is assigned a branch ID that links directly to the mutational analysis present in Extended Data Table 3 (Mutational analyses). Here, we identified 25 mutations, of which 9 (36%) exhibited ADAR-like signatures, consistent with ebolavirus transmission into the human population. Five mutations resulted in non-synonymous amino acid substitutions. We have included the following paragraph:

***Lines 162-166*:** Mutational analysis identified 25 mutations in total within the 2026 BDBV genome cluster, of which 9/25 (36%) showed patterns consistent with Adenosine Deaminase Acting on RNA (ADAR)-mediated editing. ADAR mutations are a signature of ebolavirus transmission into the human population. Of these 25 mutations, five resulted in non-synonymous substitutions (further detailed in Extended Data Table 3).

5. Is it feasible to estimate the timescale of the 2026 BVD outbreak? How strong is the temporal signal in the sequence dataset to allow for reliable molecular clock analysis?

Response 5: Given the number of BDBV genomes currently available from the initial cases we intend to describe here; a molecular clock analysis would not provide a robust estimate of timing of the most recent common ancestor (MRCA). At this stage, such analysis would likely introduce additional uncertainty and may create confusion regarding the emergence of the outbreak. However, we fully recognize that estimating the timescale of the outbreak is an important objective, and advocate for this in the text by including this statement: “Decentralized sequencing capacity will enable the generation of additional genomes needed to accurately estimate the timing of outbreak emergence.”

***Lines 180-184*:** Decentralized sequencing capacity will enable the generation of additional genomes to accurately estimate the date of the most recent common ancestor of the 2026 BVD outbreak.

6. Map Representation (Figure 1, Top Left Panel): The variable size of location names on the map, such as Kampala appearing much larger than Wakiso, seems inconsistent. Standardizing the size of location names would improve readability and presentation quality.

Response 6: The variable size of the location labels was intentionally used to emphasize the three hotspot locations associated with the known BVD outbreaks: 2007 outbreak in Bundibugyo district, Uganda; 2012 outbreak in Isiro province, DRC; and the 2026 outbreak in Ituri province, DRC. We agree that Kampala should not have been displayed at the same size as these outbreak-associated locations and have revised the figure accordingly.

7. Additionally, it would be useful to use consistent coloring for the map dots, aligned with the map's color legend on the right for clarity and interpretability.

Response 7: To further improve readability and interpretability of Figure 1, we have harmonized the color scheme across all panels. Specifically, the colors of the tips in the phylogenetic trees are now consistent with those used on the map (dark green). In addition, we have revised the tip labels to clearly indicate the locations and dates, (e.g., PP_006Y8NC|Bunia|DRC|2026-05-15). We truly appreciate this comment and believe that this change indeed significantly improves the clarity and interpretation of the figure. It also allowed us to use additional color annotations to resolve the ADAR-like mutational pattern.

On behalf of all authors,

Yours sincerely,

Tony Wawina-Bokalanga.

MD, PhD

Institut National de Recherche Biomédicale (INRB)
Department of Epidemiology and Global Health
Kinshasa
Democratic Republic of the Congo

July 24, 2026

Reviewer #1:

I am satisfied with the edits made. A couple of minor additional comments:

- The introduction is still very acronym-heavy.

Response We have revised the introduction to reduce the number of acronyms. However, acronyms such as "BDBV" for Bundibugyo virus and "BVD" for Bundibugyo virus disease have been retained where appropriate to improve readability and avoid repeated use of the full terms throughout the manuscript.

- GenBank accession NC_014373 being used as the reference genome on which to call mutations. Would be helpful (and important) to add a sentence as to why this one was chosen as the reference genome.

Response We have added a sentence clarifying the rationale for the reference genome selection and revised the text as follows:

***Lines 350-552*:** The reference genome was selected based on two criteria: (i) the availability of a complete genome sequence from the 2012 Bundibugyo outbreak, and (ii) its close genetic similarity to the initial dataset, making KC545393.1 the appropriate reference genome.

- In the Table providing the newly sequenced genomes, under the column GC content, sometimes the percentage is specified as XX.X% and sometimes XX,X%. Standardize.

Response Corrected

Reviewer #2:

The authors provided a comprehensive response to my queries.

Response Thank you very much.

Reviewer #3:

The authors have addressed my previous comments, and the revised manuscript appears to be substantially improved. I have only a few minor suggestions for further refinement.

Response We thank the reviewer for this feedback.

1. While Figure 1 has been considerably improved, particularly in terms of colour consistency, I would further suggest ensuring that the symbol shapes in the right-hand tree are fully consistent with those used on the map. At present, taxa in the right-hand tree are shown as circles (DRC) and diamond (Uganda); however, the labels of samples from Uganda on the map notation should also be represented as diamonds to match the tree taxa shape.

Response 1. Thank you. We have revised Figure 1 so that the symbols used for the Uganda samples on the map are represented as diamonds, ensuring consistency with the symbol shapes used in the corresponding phylogenetic tree (see attached figure 1 below).

2. Although the authors justify the collapsed BDBV tree as a clearer way to highlight relationships among outbreaks and to support the proposed novel zoonotic spillover event, I still recommend including the complete, uncollapsed tree in the supplementary materials for readers who are interested in the full phylogenetic context. In that supplementary full tree, the authors could consider using different branch-length scales for within-outbreak and between-outbreak clades to improve readability without losing either resolution or overall phylogenetic structure.

Response 2. We have added the complete, uncollapsed phylogenetic tree as Extended Data Fig. 1 (line 151) to provide full phylogenetic context and allow readers to examine genome relationships in greater detail.

Extended Data Fig. 1: Uncollapsed phylogenetic tree of Bundibugyo virus genomes showing the complete phylogenetic context.

3. I also suggest adding a "#" symbol before the node labels to distinguish them more clearly from bootstrap values and thereby avoid potential confusion.

Response 3. We have incorporated a '#' symbol before the node labels throughout the figure to clearly distinguish them from bootstrap values and improve figure readability (Figure 1).

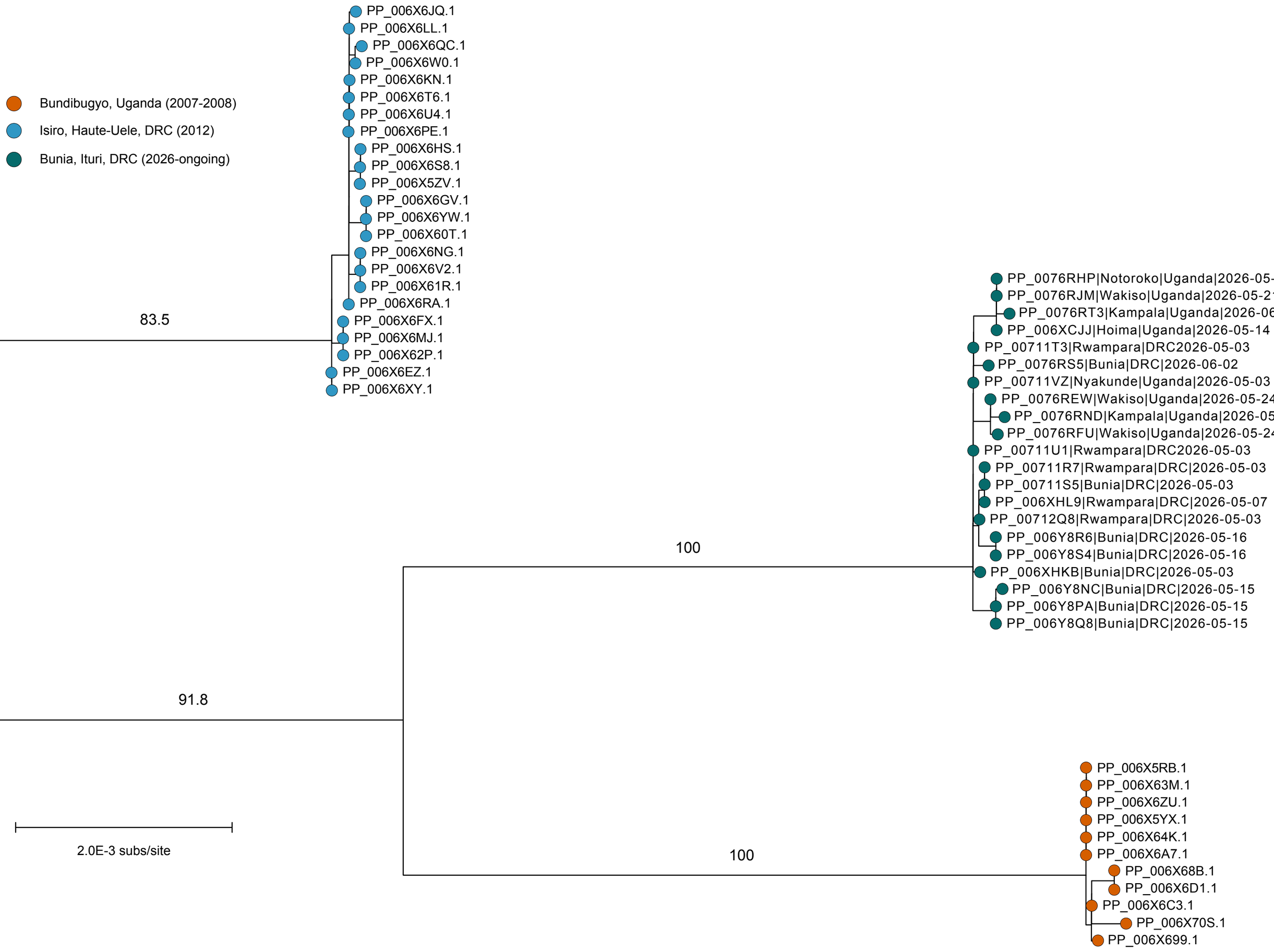
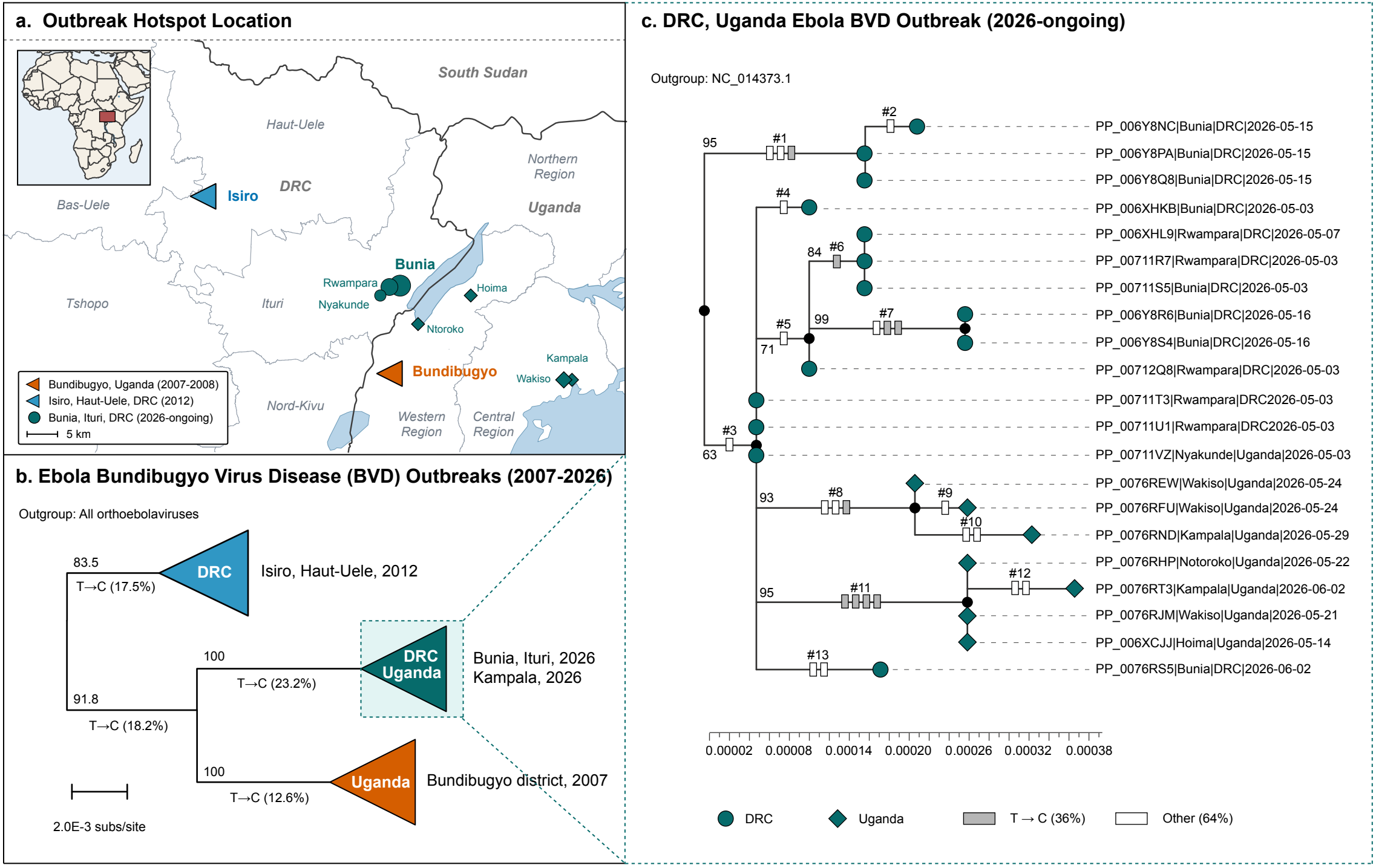

Bundibugyo, Uganda (2007-2008)
Isiro, Haute-Uele, DRC (2012)
Bunia, Ituri, DRC (2026-ongoing)
PP_006X6JQ.1
PP_006X6LL.1
PP_006X6QC.1
PP_006X6W0.1
PP_006X6KN.1
PP_006X6T6.1
PP_006X6U4.1
PP_006X6PE.1
PP_006X6HS.1
PP_006X6S8.1
PP_006X5ZV.1
PP_006X6GV.1
PP_006X6YW.1
PP_006X60T.1
PP_006X6NG.1
PP_006X6V2.1
PP_006X61R.1
PP_006X6RA.1
PP_006X6FX.1
PP_006X6MJ.1
PP_006X62P.1
PP_006X6EZ.1
PP_006X6XY.1
83.5
PP_0076RHP|Notoroko|Uganda|2026-05-22
PP_0076RJM|Wakiso|Uganda|2026-05-21
PP_0076RT3|Kampala|Uganda|2026-06-02
PP_006XCJJ|Hoima|Uganda|2026-05-14
PP_00711T3|Rwampara|DRC2026-05-03
PP_0076RS5|Bunia|DRC|2026-06-02
PP_00711VZ|Nyakunde|Uganda|2026-05-03
PP_0076REW|Wakiso|Uganda|2026-05-24
PP_0076RND|Kampala|Uganda|2026-05-29
PP_0076RFU|Wakiso|Uganda|2026-05-24
PP_00711U1|Rwampara|DRC2026-05-03
PP_00711R7|Rwampara|DRC|2026-05-03
PP_00711S5|Bunia|DRC|2026-05-03
PP_006XHL9|Rwampara|DRC|2026-05-07
PP_00712Q8|Rwampara|DRC|2026-05-03
PP_006Y8R6|Bunia|DRC|2026-05-16
PP_006Y8S4|Bunia|DRC|2026-05-16
PP_006XHKB|Bunia|DRC|2026-05-03
PP_006Y8NC|Bunia|DRC|2026-05-15
PP_006Y8PA|Bunia|DRC|2026-05-15
PP_006Y8Q8|Bunia|DRC|2026-05-15
100
91.8
100
2.0E-3 subs/site
PP_006X5RB.1
PP_006X63M.1
PP_006X6ZU.1
PP_006X5YX.1
PP_006X64K.1
PP_006X6A7.1
PP_006X68B.1
PP_006X6D1.1
PP_006X6C3.1
PP_006X70S.1
PP_006X699.1

Fig. 1 : **Geographic distribution and phylogenetic analyses of Bundibugyo virus (BDBV) genomes. Upper left panel:** Map of eastern DRC and Uganda showing all known BVD outbreak hotspot locations. **Lower left panel:** Maximum-likelihood phylogeny of all known BVD outbreaks inferred using the +gamma model in IQTree2. Ultrafast bootstrap support values are shown for each branch. The scale is given in units of substitutions per site. The current outbreak is highlighted in dark green, previous outbreaks in orange (Uganda, 2007) and blue (DRC, 2012). **Right panel:** Maximum-likelihood phylogeny of 21 high-quality 2026 BDBV genomes using the same parameters as described above. Mutations identified on each branch are indicated (ADAR-like and other), and branch IDs (i.e., 1-13) correspond to those listed in Extended Data Table 3 (Mutational analysis), providing a high-resolution view of the evolutionary relationships and mutations acquired during the 2026 outbreak. PP_006Y8R6 and PP_006Y8S4 represent two samples collected from the same patient.

On behalf of all authors,
Yours sincerely,

Tony Wawina-Bokalanga.
MD, PhD


a. Outbreak Hotspot Location
South Sudan
Haut-Uele
Northern Region
DRC
Isiro
Bas-Uele
Uganda
Bunia
Rwampara
Nyakunde
Hoima
Tshopo
Ituri
Ntoroko
Kampala
Bundibugyo
Wakiso
Nord-Kivu
Western Region
Central Region
Bundibugyo, Uganda (2007-2008)
Isiro, Haut-Uele, DRC (2012)
Bunia, Ituri, DRC (2026-ongoing)
5 km
b. Ebola Bundibugyo Virus Disease (BVD) Outbreaks (2007-2026)
Outgroup: All orthoebolaviruses
83.5
DRC
T→C (17.5%)
Isiro, Haut-Uele, 2012
100
DRC Uganda
T→C (23.2%)
Bunia, Ituri, 2026
Kampala, 2026
91.8
T→C (18.2%)
100
Uganda
T→C (12.6%)
Bundibugyo district, 2007
2.0E-3 subs/site
c. DRC, Uganda Ebola BVD Outbreak (2026-ongoing)
Outgroup: NC_014373.1
PP_006Y8NC|Bunia|DRC|2026-05-15
PP_006Y8PA|Bunia|DRC|2026-05-15
PP_006Y8Q8|Bunia|DRC|2026-05-15
PP_006XHKB|Bunia|DRC|2026-05-03
PP_006XHL9|Rwampara|DRC|2026-05-07
PP_00711R7|Rwampara|DRC|2026-05-03
PP_00711S5|Bunia|DRC|2026-05-03
PP_006Y8R6|Bunia|DRC|2026-05-16
PP_006Y8S4|Bunia|DRC|2026-05-16
PP_00712Q8|Rwampara|DRC|2026-05-03
PP_00711T3|Rwampara|DRC2026-05-03
PP_00711U1|Rwampara|DRC2026-05-03
PP_00711VZ|Nyakunde|Uganda|2026-05-03
PP_0076REW|Wakiso|Uganda|2026-05-24
PP_0076RFU|Wakiso|Uganda|2026-05-24
PP_0076RND|Kampala|Uganda|2026-05-29
PP_0076RHP|Notoroko|Uganda|2026-05-22
PP_0076RT3|Kampala|Uganda|2026-06-02
PP_0076RJM|Wakiso|Uganda|2026-05-21
PP_006XCJJ|Hoima|Uganda|2026-05-14
PP_0076RS5|Bunia|DRC|2026-06-02
0.00002
0.00008
0.00014
0.00020
0.00026
0.00032
0.00038
DRC
Uganda
T → C (36%)
Other (64%)