# Evidence of dengue virus transmission and a diverse *Aedes* mosquito virome on the Democratic Republic of Congo-Angola border

**Running title: *Aedes* mosquito virome in the DRC**

Wenqiao He[1#], Thierry Bobanga[2], Anne Piantadosi[3,4], Zachary R. Popkin-Hall[1], Fabien Vulu[2], Matthew H. Collins[4], Melchior M. Kashamuka[5], Antoinette K. Tshefu[5], Jonathan J. Juliano[1*], Jonathan B. Parr[1*#]

[1] Institute for Global Health and Infectious Diseases and Division of Infectious Diseases, Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2] Department of Tropical Medicine, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of the Congo

[3] Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, GA, USA

[4] Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, GA, USA

[5] Kinshasa School of Public Health, Kinshasa, Democratic Republic of the Congo

* Co-senior authors

[#] Corresponding authors: wenqiao@email.unc.edu; jonathan_parr@med.unc.edu

## Abstract

*Aedes* mosquitoes are widely distributed across the Democratic Republic of Congo (DRC), and are major vectors of dengue (DENV), Zika, chikungunya (CHIKV), and yellow fever (YFV) viruses. While the high burden of malaria in the DRC receives considerable attention, arboviruses remain understudied. In the setting of recent CHIKV and YFV outbreaks in southwestern DRC, we collected *Aedes* mosquitoes in three areas of Kimpese, DRC, near the Angola border, to investigate their virome. Metagenomic and targeted sequencing of eight randomly selected field mosquito pools of 115 mosquitoes confirmed high-confidence DENV reads and human blood meals in six (75%) and eight (100%) pools, respectively. We find diverse mosquito viromes including other known and putative human and animal viruses. Our findings provide strong evidence of endemic DENV transmission along the DRC-Angola border and illustrate the potential of wild-caught mosquitoes for xenosurveillance of emerging pathogens.

**Keywords:** *Aedes* mosquito, dengue, Democratic Republic of Congo, viral metagenomic

## Introduction

Arthropod-borne pathogens, including viruses, bacteria, and parasites, account for more than 17% of all infectious diseases [1,2], and over 700 arboviruses have been documented to infect humans across a wide geographic distribution [3,4]. The number of outbreaks and endemic infections caused by emergent and re-emergent mosquito-borne arboviruses have increased worldwide over the last two decades [5–9]. *Aedes* mosquitoes are particularly important vectors for multiple important human pathogens, including dengue virus (DENV), Zika virus (ZIKV), yellow fever virus (YFV), and chikungunya virus (CHIKV) [10–13]. The warm climate and suitable environmental conditions in tropical regions support the proliferation of both mosquitoes and the viruses they transmit. As a result, tropical regions tend to have higher diversity and prevalence of arboviruses than other regions. However, arbovirus transmission is understudied across Africa, where improved understanding is urgently needed to guide roll out of newly available DENV vaccines, in particular [14–16].

The Democratic Republic of the Congo (DRC) is the largest country in central Africa with a population of over 100 million, and a tropical climate conducive to mosquitoes. Approximately 248 mosquito species are thought to reside in the DRC [17], including *Aedes aegypti* and *Aedes albopictus*, both of which are major disease vectors [9,18]. A recent survey found that *Ae. albopictus* is expanding its distribution in the DRC and displacing native *Aedes* species [19]. Chikungunya, YFV, and DENV outbreaks and infections are known to occur in the DRC [20–22], and increasing reports of mosquito-borne and emerging viral diseases have been recorded [23]. Considerable attention is devoted to the high burden of malaria [24,25] in the DRC, but arboviruses remain neglected, with limited studies of humans and mosquitoes to-date [3,21–23,26]. The lack of diagnostic tools and surveillance systems for these diseases hinders our understanding of their impact and epidemiology [27].

Metagenomic sequencing approaches can be used for identification of viral DNA and/or RNA within diverse sample types, and have recently been deployed for mosquito virome analyses [28–33]. To improve our understanding of the role of *Aedes* mosquito vectors in arboviral transmission in the DRC, we collect and sequence the virome of mosquito pools from three areas of Kimpese, a region near the Angola border which has experienced recent arboviral outbreaks (**Figure 1**). We observe diverse viromes and identify strong evidence of active DENV transmission in the region, corroborated by viral sequences and mosquito blood meal analysis.

## Methods

### Mosquito collection

Adult *Aedes* mosquitoes were collected from three health areas in Kimpese, DRC, Ceco (Kimpese city), Malanga, and Viaza (**Figure 2a**), using electric aspirators (Prokopack Aspirator, John W. Hock, Gainesville, USA) and BG sentinel traps (Biogents Inc, Regensburg, Germany) baited with dry ice. Mosquitoes were collected from April 23 - 30, 2022, during the rainy season when mosquito density is high. All collections were performed outdoors, and GPS coordinates along with a brief description of collection sites were recorded.

Collected mosquitoes were morphologically identified to the species level following the Huang morphological key [34] and then pooled according to sample collection sites. All mosquito pools were preserved in DNA/RNA Shield (ZYMO research, CA, USA) and stored at -80°C.

### Nucleic acid extraction

Eight field *Aedes* mosquito pools were randomly selected from the three health areas: five from Kimpese city, one from Malanga, and two from Viaza (**Figure 2b**). The selected mosquito pools were washed with 70% ethanol and rinsed twice with molecular-grade water before removal of mosquito heads [35]. Total RNA was extracted from mosquito bodies from the selected field mosquito pools, one laboratory un-infected *Ae. aegypti* mosquito pool from BEI Resources, and one water control using the Quick-RNA Tissue/Insect kit (ZYMO research, CA, USA) following the manufacturer's instructions. DNase I treatment using DNase I Set (ZYMO research, CA, USA) was carried out during the RNA extraction process.

### Mosquito species confirmation

To confirm the mosquito species, two regions within the cytochrome c oxidase I (*COI*), and one region within the second internal transcribed spacer (*ITS2*) genes were amplified from RNA extracted from the mosquito pools using published primers [36–38]. Amplicon libraries were constructed using the Native Barcoding Kit 96 V14 and sequenced on a Flongle Flow Cell (R10.4.1) using the MinION platform (Oxford Nanopore Technologies, Oxford, United Kingdom).

### Ribosomal RNA depletion, library preparation, and viral metagenomic sequencing

To increase sequencing efficiency, mosquito ribosomal RNA (rRNA) depletion was carried out before library preparation using the NEBNext RNA Depletion Core Reagent set (New England BioLabs, MA, USA) with published probes targeting *Aedes* mosquitoes [39]. Sequencing libraries were prepared and barcoded using the NEBNext Ultra II Directional RNA Library Prep Kit and the Multiplex Oligos for Illumina® (96 Unique Dual Index Primer Pairs) kit (New England BioLabs, Massachusetts, USA), respectively. The barcoded libraries were sequenced using the NovaSeq S4 platform (Illumina, California, USA) at the High-Throughput Sequencing Facility (HTSF) at the University of North Carolina at Chapel Hill to generate 150 bp paired-end reads.

**Dengue virus confirmation by real-time PCR**

Complementary DNA (cDNA) was synthesized from RNA extracted from field mosquito pools using the Transcriptor First Strand cDNA Synthesis Kit (Roche, IN, USA). We then used a published real-time PCR assay to detect DENV cDNA [40].

**Mosquito blood meal investigation**

The 16S ribosomal RNA (*16S*) and cytochrome b (*cytB*) genes were amplified from RNA extracted from field mosquito pools to determine mosquito blood meal sources using published primers [41,42]. Amplicon libraries were constructed using Native Barcoding Kit 96 V14 and sequenced on a Flongle Flow Cell using the MinION platform as described above.

**Data analysis**

To investigate viruses in the mosquito pools, Illumina NovaSeq sequencing reads were demultiplexed into individual samples based on their barcode sequences. Paired-end reads were then merged into single reads based on overlap detection using `BBMerge` [43], followed by adapter removal using `Trimmomatic` [44]. Reads mapping to human (GRCh38.p14), *Ae. albopictus* (Aalbo_primary.1), and *Ae. aegypti* (AaegL5.0) genomes, as well as the *Ae. simpsoni* CO1 gene (KT881399.1) were removed using `bwa-mem2` [45] and `SAMtools` [46].

Reads remaining in the laboratory un-infected *Ae. aegypti* mosquito pool and water control pool were assembled into contigs using `metaSPAdes` [47]. These contigs were considered as contaminating contigs. Reads in field mosquito pools were aligned to the contaminating contigs using `bwa-mem2`, and mapping reads were removed using `SAMtools`. Taxonomic

classification of the remaining reads in field mosquito pools were performed using KrakenUniq with its default mode and standard NCBI-nt and viral databases [48].The relative abundance of viral genera and families was calculated based on the KrakenUniq reports using the default nt database. This was done by dividing the number of reads mapped to specific viruses or taxonomic groups by the total number of viral reads, then multiplying the result by 100. Reads classified as viruses were extracted for further analysis.

Principal component analysis (PCA) was performed based on the viral genus annotation. Blast (version 2.14.1) was used to confirm the classification of reads mapping to human and animal pathogens and potential pathogens. Blast-confirmed reads mapping to human and animal pathogens or potential pathogens were assembled into contigs using SPAdes (version 4.0.0). The longest contigs or merged paired-end reads for each virus in the detected pools were selected and aligned with reference sequences of different genotypes from different hosts/regions from the GenBank database using the ClustalW multiple sequence alignment program in MEGA (version 7.0; Oxford Molecular Ltd., UK). Phylogenetic trees were generated using the neighbor-joining (NJ) method in MEGA, with 1,000 bootstrap replicates.

For Nanopore sequencing data used for mosquito species determination and blood meal analysis, raw data in Fast5 file format from the sequencer were used for base-calling with Guppy (version 5.6.7), with a phred quality score ≥ 6 (Q score ≥ 6). Reads were demultiplexed according to their barcodes, which were then trimmed. BLAST was used to investigate the mosquito species and blood sources (calls were made for hits with e-value < 1e-6 and length > 100bp).

All statistical analyses were performed and visualized in R software (version 4.2.0; R Core Team, Vienna, Austria) in RStudio (version 2022.02.2) with *ggrepel* (version 0.9.5), *ggplot2* (version 3.5.1), *tidyverse* (version 2.0.0), *eulerr* (version 7.0.2), and *dplyr* (version 1.1.4) packages. Mapping was done using *sf* (version 1.0.16), *rnaturalearth* (version 1.0.1), and *rnaturalearthdata* (version 1.0.0) packages in R software. Data for mapping was downloaded from the Global Administrative Areas Database (GADM) [49] and the Humanitarian Data Exchange (HDX) [50].

## Results

### Mosquito collection and species confirmation

More than 600 adult *Aedes* mosquitoes were collected and distributed into 37 pools based on sample collection sites (21 pools from Kimpese city, 6 pools from Malanga, and 10 pools from Viaza). Three species of *Aedes* mosquitoes were identified by morphology: *Aedes albopictus (n=653, 98.34%)*, *Aedes aegypti (n=8, 1.20%)*, and *Aedes simpsoni (n=3, 0.45%)*. Among the eight sequenced pools, all contained majority *Ae. albopictus* (**Table S1**) based on their *ITS2* and *COI* gene sequences.

**Sequencing output**

A total of 3,795,246,904 raw reads were obtained by Illumina sequencing. A range of 78.2-84.4% paired-end reads in each field mosquito pool were successfully merged. After filtering out low quality reads and those mapping to host and contaminating contigs (**Table S2**), 185,621,738 reads remained, of which 144,469,836 were unclassified reads (66.2% of total) that did not map to any sequences in the Krakenuniq default nt database. Adapter content was not found in unclassified reads. Classified reads were mapped to known viruses, bacteria, eukaryota, and archaea (**Table S3**).

**Viral metagenomic analysis at the family and genus levels**

Results confirm a highly diverse virome in wild-caught *Aedes* from the DRC, with viruses specific to invertebrates, vertebrates, plants, bacteria, fungi, and protozoa, as well as an unclassified group of viruses.

Based on alignment of the NCBI-nt database, we identify 51 viral families in the field mosquito pools, ranging from 7-29 viral families per pool. A total of 17 shared viral families (**Figure 3a** and **Supplementary Material**) are identified in mosquitoes across all three sample collection areas, with reads mapping to *Solemoviridae* and *Flaviviridae* detected in all field mosquito pools. The viral family *Solemoviridae* has the highest relative abundance in most of the field mosquito pools collected in Kimpese city and Viaza, while *Permutotetraviridae* has the highest relative abundance in mosquitoes collected in Malanga (**Figure 3b**). Except for the mosquito pool from Malanga, *Flaviviridae* is among the top three viral families in other mosquito pools (**Figure 3b**). High relative abundance of *Picornaviridae* and *Dicistroviridae* is also observed in one mosquito pool collected in the Kimpese city (K2).

We observe a range of 5 to 24 viral genera per field mosquito pool, comprising 64 total viral genera. Six viral genera are present in field mosquito pools collected from all sample collection areas (**Figure 3c** and **Supplementary Material**), with *Orthoflavivirus* found in all field mosquito pools tested. *Orthoflavivirus*, which comprises viruses formerly classified as

*Orthoflavivirus* and *Flavivirus* genera[51], is the most abundant viral genus in most mosquito pools collected from Kimpese city and Viaza, while the insect virus *Jonvirus* is the genus with the highest relative abundance in the mosquito pool collected in Malanga (**Figure 3d**). PCA based on viral genus annotation to explore clustering by sample collection area is shown in **Figure S1**.

**Insect-specific viruses**

We also find a large number of insect-specific viruses in field mosquito pools, with most belonging to the families *Solemoviridae* and *Flaviviridae*. *Wenzhou sobemo-like virus 4*, *Hubei mosquito virus 2*, *Guangzhou sobemo-like virus*, *Sichuan mosquito sobemo-like virus*, and *Aedes flavivirus* are the most prevalent species (**Table S4**). Except for *Wenzhou sobemo-like virus 4* (6/8, 75%), the other four insect-specific viruses are detected in 7 out of 8 (87.5%) mosquito pools with high relative abundance.

**Phylogenetic analysis of reads mapping to pathogens and potential pathogens**

Blast-confirmed reads mapping to viral species related to human or animal diseases, including DENV, as well as *human pegivirus*, *human blood-associated dicistrovirus*, and *bat faecal associated dicistrovirus 4*, are also observed in the mosquito pools (**Figure 4**). Though SPAdes does not successfully assemble long contigs from DENV reads in this study, BLAST-confirmed, merged paired-end reads mapping to DENV are found in 6 out of the 8 (75%) field mosquito pools (**Table S5**). Reads mapping to DENV-2 in one mosquito pool (K5) are highly similar to one sequence from Malaysia (MH048672.1 10664-10770, 107 bp, nucleotide identity = 94%, **Figure 5a**). Reads mapping to DENV-4 found in all six positive pools are short (< 100bp) but show high similarity to the 5'-untranslated (UTR) region of a published sequence from Thailand (MG601754.1 1-90, nucleotide identity ≥ 99%) (**Figure 5b**). Phylogenetic analysis of these short, merged paired-end reads suggests the presence of DENV types 2 and 4 in the DRC, though bootstrap support is limited.

We find high similarity between the reads mapped to *human pegivirus*, *human blood-associated dicistrovirus*, and *bat fecal-associated dicistrovirus 4*, and their respective published sequences. Additional detail about these viruses is provided in **Figures S2-S4**, **Table S6**, and the **Supplementary Material**.

**PCR confirmation of DENV and mosquito blood meal investigation**

Real-time PCR targeting the DENV polyprotein gene, distinct from genomic regions included in our phylogenetic analysis, confirms DENV RNA in the three mosquito pools with highest

DENV read counts in metagenomic analysis (**Table S5**), indicating consistency between next-generation sequencing and real-time PCR within its limits of detection. Investigation of mosquito blood meals using amplicon sequencing showed that mosquitoes from all selected pools had taken blood from *Homo sapiens* (**Figure 6**). Additionally, we detected reads mapping to non-human primates (3/8 pools), rodents (4/8 pools), and sheep (1/8 pools).

## Discussion

Our findings confirm the presence of dengue and other viruses with human or animal pathogenic potential in wild-caught *Aedes* mosquitoes circulating near the DRC's border with Angola. The presence of human blood in all mosquito pools strongly suggests active dengue transmission in the Kimpese region, including peri-urban and rural zones outside of the city. The combination of a diverse mosquito virome and the presence of blood meals from multiple mammals suggests robust opportunities for arthropod-borne viral transmission between species, as well as for spillover events. Together, our experience suggests a future role for xenosurveillance and metagenomic sequencing as a complement to existing data systems developed to monitor for emerging infectious diseases in countries like the DRC.

Multiple arboviral disease outbreaks have been documented in the DRC during the last century, including yellow fever, dengue, chikungunya, and others [3,21,22,26,52]. Except for DENV, we did not find other known pathogenic arboviruses related to human and vertebrate diseases in the selected *Aedes* pools. Detection of DENV-2 and DENV-4 in the mosquito pools suggests active DENV circulation in the Kimpese region. Three of the four DENV serotypes [53] have been recorded in the human population of the DRC (DENV-1, DENV-2, and DENV-3), with DENV-1 being the most frequently reported [3]. One study in Singapore reported that the majority of the dengue hemorrhagic fever cases involving DENV-2 were from the cosmopolitan genotype [54]. DENV-2 detected in our study is most similar to cosmopolitan genotype of DENV-2, which has been previously reported in the DRC [55]. However, the DENV-2 reads detected in our study show closer similarity to cosmopolitan DENV-2 from Malaysia than those previously reported from the DRC. DENV-4 was thought to have had a more limited geographic distribution than other serotypes [56]. However, our study provides the first evidence of DENV-4 in *Aedes* mosquitoes from the DRC, suggesting this serotype may also infect people in this region.

Data from all *Aedes* mosquito pools in our study confirm human blood meals. Previous studies have reported high dietary diversity of *Aedes* mosquitoes, with a predominant preference for mammalian hosts [57–59]. We corroborate these findings, with diverse mammalian blood meals observed across mosquito pools, including blood from non-human

260 primates. Non-human primates can be amplifying hosts and play considerable roles in the
261 sylvatic cycle of DENV transmission, which can spark DENV epidemics in nearby human
262 populations [60]. Detection of rodent and sheep genomes in the collected mosquito pools
263 illustrate the diversity of hosts of *Aedes*, though further research is needed to confirm the
264 roles of these hosts in DENV transmission, including the contribution to human infection.
265 Notably, three (K2, K4, and V3) mosquito pools appear to have exclusively fed on human
266 blood. The presence of DENV RNA in two of these pools (K4 and V3) strongly suggests
267 active DENV transmission in these communities.

268 In addition to DENV, we identify reads mapping to other potential human and animal
269 pathogens, including *human pegivirus*, *bat faecal associated dicistrovirus 4*, and *human*
270 *blood-associated dicistrovirus*. Detection of these non-mosquito borne viruses likely stems
271 from mosquito blood meals, and highlights the possibility of using field-caught *Aedes* for
272 xenosurveillance of multiple pathogens. Human pegivirus (HPgV-1, GBV-C) is a
273 below-species classification of *Pegivirus platyrrhini* within the *Pegivirus* genus of the
274 *Flaviviriade* family with seven genotypes. A previous study reported presence of pegivirus
275 genotypes 1 and 5 in the DRC [61]. We report one polyprotein gene sequence highly similar to
276 that of human pegivirus genotype 2 (MZ099567.1), indicating circulation of human pegivirus
277 genotype 2 in the DRC. The pathogenicity of human pegivirus is unclear, but a survival
278 benefit among patients who are coinfected with HIV has been suggested [62]. We also report
279 the first detection of *human blood-associated dicistrovirus* and *bat faecal associated*
280 *dicistrovirus 4* in the DRC. Dicistrovirus has been recently reported in blood samples from
281 febrile patients in Tanzania and Nigeria [63,64], the intestinal content of a captive red squirrel
282 with enteritis [65], and has been identified in free-living gorillas in neighboring Republic of
283 Congo [66]. Interestingly, we found *bat faecal associated dicistrovirus 4* reads in a mosquito
284 pool (K2) that exclusively fed on human blood, suggesting that it came from an infected
285 human. However, the pathogenicity of these viruses are unclear.

286 Our findings also shed light on circulating *Aedes* species along the DRC-Angola border.
287 Kimpese comprises a small city and surrounding region where more than 100,000 people
288 are thought to live [67], though the DRC has not performed a formal census since the 1950s. It
289 is located in the Kongo Central province, has a tropical climate, and spans mountains,
290 valleys, and grasslands. *Aedes aegypti* and *Ae. albopictus* are the main vectors of several
291 major pathogenic arboviruses, with *Ae. aegypti* being more competent for DENV and ZIKV
292 transmission [68]. *Ae. aegypti* was previously the dominant *Aedes* species in the DRC, until
293 recent, rapid replacement of native *Aedes* species by *Ae. albopictus* in the western and
294 northern regions of DRC [19]. Our findings corroborate these changes in the DRC *Aedes*

population; most collected *Aedes sp.* mosquitoes are *Ae. albopictus* (98.3%). The highly diverse viromes in our *Aedes* mosquito pools include viruses specific to invertebrates, vertebrates, plants, bacteria, fungi, and protozoa. These reads likely reflect the mosquitoes' ingestion of blood, nectar, and environment substrates [69,70]. However, the large number of unclassified reads within all pools confirms how much remains to be learned about this important mosquito population.

Several limitations of our study should be highlighted. First, mosquitoes were not pooled based on morphological mosquito species. Second, our metagenomic analysis focuses on *Aedes*. Analysis of other mosquito genera would provide broader insights into differences in arboviral transmission and viromes in the DRC. Third, despite numerous reads mapping to DENV, we did not have sufficient genomic coverage to construct whole-genomes. Nor can metagenomic sequencing tell us whether infectious virions are present or definitively determine dengue types in circulation. Fourth, the use of mosquito pools prevents us from estimating the prevalence of viruses in individual mosquitoes. However, pooling improved the yield of our sequencing efforts. Finally, viral metagenomic techniques have high sensitivity but low specificity. To overcome this shortcoming, we employ a conservative approach leveraging multiple bioinformatic tools and complemented by real-time PCR, both of which confirm the presence of DENV in our sample set.

In conclusion, our study provides new insights into viruses infecting *Aedes* in the DRC and demonstrates the potential of wild-caught mosquito xenosurveillance. We use virome and blood-meal analysis to improve understanding of *Aedes* mosquito feeding behavior in the DRC, and present evidence of endemic DENV transmission in the Kimpese region near the DRC's border with Angola DRC. These findings are particularly relevant to ongoing discussions about where to offer newly approved dengue vaccines and indicate a need for larger human studies to define dengue epidemiology in the DRC and across Africa [71]. They also suggest a role for sustained vector control efforts targeting *Aedes* in the DRC, extending beyond outbreak response, and highlight opportunities for future arthropod-borne disease studies in the region.

## Acknowledgements

We thank Daniel Neafsey for sharing mosquito blood meal sequencing protocols and advice. The following reagent was provided by the NIH/NIAID Filariasis Research Reagent

Resource Center for distribution through BEI Resources, NIAID, NIH: Uninfected *Aedes aegypti*, Strain Black Eye Liverpool (Frozen), NR-48920.

The authors used an artificial intelligence language model for English language editing during the writing process. However, the manuscript is original to the authors, who take responsibility for its content.

## Author contributions

W.H., J.J.J, and J.B.P. conceptualized the study and designed experiments. T.B. and F.V. collected field mosquito samples. M.M.K. and A.K.T. helped coordinate these collections. W.H. performed laboratory tests. W.H., A.P., Z.R.P.-H., and M.H.C. performed data analyses. W.H. prepared the figures and wrote the first draft. All authors reviewed and approved the final draft.

## Funding

Mosquito collection was funded by a Yang Biomedical Scholar award from University of North Carolina (JJJ). This work was supported in part by NIAID K24AI134990 (JJJ), funds from the UNC Office for Research (JBP), and the Bill & Melinda Gates Foundation (INV-050353 to JBP). Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has already been assigned to the Author Accepted Manuscript version that might arise from this submission.

## Competing interests:

JBP reports research support from Gilead Sciences, non-financial support from Abbott Laboratories, and past consulting for Zymeron Corporation, all outside the scope of the current manuscript.

## Data and code availability

NovaSeq and MinION sequencing data presented in the study will be deposited in the NCBI Sequence Read Archive (BioProject ID: pending). Analysis code will be made publicly available at https://github.com/orgs/IDEELResearch/DRCAedesSeq.

## References


1. World Health Organization. Vector-borne diseases. WHO: Geneva, 2024. Available at https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases [Accessed 18 December 2024].
2. Kading, R. C., Brault, A. C. & Beckham, J. D. Global Perspectives on Arbovirus Outbreaks: A 2020 Snapshot. *Trop Med Infect Dis* **5**, (2020).
3. Mbanzulu, K. M. *et al.* Mosquito-borne viral diseases in the Democratic Republic of the Congo: a review. *Parasit. Vectors* **13**, 103 (2020).
4. Burt, F. J. *et al.* Chikungunya virus: an update on the biology and pathogenesis of this emerging pathogen. *Lancet Infect. Dis.* **17**, e107–e117 (2017).
5. Bharati, N. *et al.* Circulating serotypes and genotypes of dengue virus during the 2023 outbreak in Eastern Nepal. *J. Clin. Virol.* **174**, 105721 (2024).
6. Tajudeen, Y. A., Oladunjoye, I. O., Mustapha, M. O., Mustapha, S. T. & Ajide-Bamigboye, N. T. Tackling the global health threat of arboviruses: An appraisal of the three holistic approaches to health. *Health Promot Perspect* **11**, 371–381 (2021).
7. Soto-Garita, C. *et al.* Epidemiological, virological and clinical characterization of a Dengue/Zika outbreak in the Caribbean region of Costa Rica 2017-2018. *Front. Cell. Infect. Microbiol.* **14**, 1421744 (2024).
8. Chala, B. & Hamde, F. Emerging and re-emerging vector-borne infectious diseases and the challenges for control: A review. *Front. Public Health* **9**, 715759 (2021).
9. Bobanga, T., Moyo, M., Vulu, F. & Irish, S. R. First Report of *Aedes albopictus* (Diptera: Culicidae) in the Democratic Republic of Congo. *Afr. Entomol.* **26**, 234–236 (2018).
10. Poungou, N. *et al.* Mosquito-Borne Arboviruses Occurrence and Distribution in the Last Three Decades in Central Africa: A Systematic Literature Review. *Microorganisms* **12**,

(2023).

11. Janjoter, S., Kataria, D., Yadav, M., Dahiya, N. & Sehrawat, N. Transovarial transmission of mosquito-borne viruses: a systematic review. *Front. Cell. Infect. Microbiol.* **13**, 1304938 (2023).
12. Renard, A. *et al.* Interaction of Human Behavioral Factors Shapes the Transmission of Arboviruses by Aedes and Culex Mosquitoes. *Pathogens* **12**, (2023).
13. Kraemer, M. U. *et al.* The global distribution of the arbovirus vectors and. *Elife* **4**, (2015).
14. Mercy, K. *et al.* The looming threat of dengue fever: The Africa context. *Open Forum Infect. Dis.* **11**, ofae362 (2024).
15. Shikanda, H. WHO prequalifies new dengue vaccine, but Africa still not in the picture. *Nation.* May 17, 2024. Available at: https://nation.africa/kenya/health/who-prequalifies-new-dengue-vaccine-but-africa-still-not-in-the-picture-4626442 [Accessed 18 December 2024].
16. Jaenisch, T. *et al.* Dengue expansion in Africa-not recognized or not happening? *Emerg. Infect. Dis.* **20**, (2014).
17. Foley, D. H., Rueda, L. M. & Wilkerson, R. C. Insight into global mosquito biogeography from country species records. *J. Med. Entomol.* **44**, 554–567 (2007).
18. Manzambi, E. Z. *et al.* Behavior of Adult Aedes aegypti and Aedes albopictus in Kinshasa, DRC, and the Implications for Control. *Trop. Med. Infect. Dis.* **8**, (2023).
19. Vulu, F. *et al.* Geographic expansion of the introduced Aedes albopictus and other native Aedes species in the Democratic Republic of the Congo. *Parasit. Vectors* **17**, 35 (2024).
20. Otshudiema, J. O. *et al.* Yellow Fever Outbreak - Kongo Central Province, Democratic Republic of the Congo, August 2016. *MMWR Morb. Mortal. Wkly. Rep.* **66**, 335–338 (2017).
21. Willcox, A. C. *et al.* Seroepidemiology of Dengue, Zika, and Yellow Fever Viruses among Children in the Democratic Republic of the Congo. *Am. J. Trop. Med. Hyg.* **99**, 756–763 (2018).

22. Selhorst, P. *et al.* Molecular characterization of chikungunya virus during the 2019 outbreak in the Democratic Republic of the Congo. *Emerg. Microbes Infect.* **9**, 1912–1918 (2020).

23. Mbanzulu, K. M. *et al.* Mosquito-borne viruses circulating in Kinshasa, Democratic Republic of the Congo. *Int. J. Infect. Dis.* **57**, 32–37 (2017).

24. Sendor, R. *et al.* Epidemiology of Plasmodium malariae and Plasmodium ovale spp. in Kinshasa Province, Democratic Republic of Congo. *Nat. Commun.* **14**, 6618 (2023).

25. Deutsch-Feldman, M. *et al.* What is the burden of malaria in the DRC? *J. Infect. Dis.* (2020) doi:10.1093/infdis/jiaa650.

26. Messina, J. P. *et al.* The current and future global distribution and population at risk of dengue. *Nat Microbiol* **4**, 1508–1515 (2019).

27. Kajeguka, D. C. *et al.* Prevalence of dengue and chikungunya virus infections in north-eastern Tanzania: a cross sectional study among participants presenting with malaria-like symptoms. *BMC Infect. Dis.* **16**, 183 (2016).

28. Bernadus, J. B. B. *et al.* Metagenomic Insight into the Microbiome and Virome Associated with Aedes aegypti Mosquitoes in Manado (North Sulawesi, Indonesia). *Infect. Dis. Rep.* **15**, 549–563 (2023).

29. Gangopadhayya, A. *et al.* Metagenomic Analysis of Viromes of Aedes Mosquitoes across India. *Viruses* **16**, (2024).

30. Xia, H. *et al.* Comparative Metagenomic Profiling of Viromes Associated with Four Common Mosquito Species in China. *Virol. Sin.* **33**, 59–66 (2018).

31. Pan, Y.-F. *et al.* Metagenomic analysis of individual mosquito viromes reveals the geographical patterns and drivers of viral diversity. *Nat Ecol Evol* **8**, 947–959 (2024).

32. Wang, Y. *et al.* Metagenomic sequencing reveals viral diversity of mosquitoes from Shandong Province, China. *Microbiol Spectr* e0393223 (2024).

33. Oguzie, J. U. *et al.* Metagenomic sequencing characterizes a wide diversity of viruses in field mosquito samples in Nigeria. *Sci. Rep.* **12**, 7616 (2022).

34. Huang, Y.-M. *The Subgenus Stegomyia of Aedes in the Afrotropical Region with Keys to*

436 *the Species (Diptera: Culicidae)*. (Magnolia Press, 2004).

437 35. Phalen, H. *et al.* A Mosquito Pick-and-Place System for PfSPZ-based Malaria Vaccine
438 Production. *IEEE Trans. Autom. Sci. Eng.* **18**, 299–310 (2020).

439 36. Delgado-Serra, S. *et al.* Molecular Characterization of Mosquito Diversity in the Balearic
440 Islands. *J. Med. Entomol.* **58**, 608–615 (2021).

441 37. Zhang, D. X. & Hewitt, G. M. Nuclear integrations: challenges for mitochondrial DNA
442 markers. *Trends Ecol. Evol.* **11**, 247–251 (1996).

443 38. Batovska, J., Blacket, M. J., Brown, K. & Lynch, S. E. Molecular identification of
444 mosquitoes (Diptera: Culicidae) in southeastern Australia. *Ecol. Evol.* **6**, 3001–3011
445 (2016).

446 39. Fauver, J. R. *et al.* A reverse-transcription/RNase H based protocol for depletion of
447 mosquito ribosomal RNA facilitates viral intrahost evolution analysis, transcriptomics
448 and pathogen discovery. *Virology* **528**, 181–197 (2018).

449 40. Waggoner, J. J. *et al.* Single-Reaction Multiplex Reverse Transcription PCR for
450 Detection of Zika, Chikungunya, and Dengue Viruses. *Emerg. Infect. Dis.* **22**,
451 1295–1297 (2016).

452 41. Meyer, R., Höfelein, C., Lüthy, J. & Candrian, U. Polymerase chain reaction-restriction
453 fragment length polymorphism analysis: a simple method for species identification in
454 food. *J. AOAC Int.* **78**, 1542–1551 (1995).

455 42. Taylor, P. G. Reproducibility of ancient DNA sequences from extinct Pleistocene fauna.
456 *Mol. Biol. Evol.* **13**, 283–285 (1996).

457 43. Bushnell, B., Rood, J. & Singer, E. BBMerge - Accurate paired shotgun read merging
458 via overlap. *PLoS One* **12**, e0185056 (2017).

459 44. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina
460 sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

461 45. Vasimuddin, M., Misra, S., Li, H. & Aluru, S. Efficient Architecture-Aware Acceleration of
462 BWA-MEM for Multicore Systems. in *2019 IEEE International Parallel and Distributed*
463 *Processing Symposium (IPDPS)* 314–324 (IEEE, 2019).

46. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

47. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).

48. Breitwieser, F. P., Baker, D. N. & Salzberg, S. L. KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. *Genome Biol.* **19**, 198 (2018).

49. Database of Global Administrative Areas (GADM) (2024). *University of California, Berkeley*. Available online: http://www.gadm.org [18 December, 2024].

50. United Nations Office for Coordination of Humanitarian Affairs. Humanitarian Data Exchange (v1.86.1). Available online: https://data.humdata.org/ [18 December, 2024].

51. Postler, T. S. *et al.* Renaming of the genus Flavivirus to Orthoflavivirus and extension of binomial species names within the family Flaviviridae. *Arch. Virol.* **168**, (2023).

52. Kraemer, M. U. G. *et al.* Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015-16: a modelling study. *Lancet Infect. Dis.* **17**, 330–338 (2017).

53. Rahman, A. *et al.* Unveiling ancestral relations, host–pathogen interactions and comparative viral miRNA insights of dengue virus serotypes. *Network Modeling Analysis in Health Informatics and Bioinformatics* **10**, 27 (2021).

54. Yung, C.-F. *et al.* Dengue serotype-specific differences in clinical manifestation, laboratory parameters and risk of severe disease in adults, singapore. *Am. J. Trop. Med. Hyg.* **92**, 999–1005 (2015).

55. Selhorst, P. *et al.* Phylogeographic analysis of dengue virus serotype 1 and cosmopolitan serotype 2 in Africa. *Int. J. Infect. Dis.* **133**, 46–52 (2023).

56. Phadungsombat, J., Nakayama, E. E. & Shioda, T. Unraveling dengue virus diversity in Asia: An epidemiological study through genetic sequences and phylogenetic analysis. *Viruses* **16**, 1046 (2024).

57. Garamszegi, L. Z. Host diversity of Aedes albopictus in relation to invasion history: a meta-analysis of blood-feeding studies. *Parasit. Vectors* **17**, 411 (2024).

58. Richards, S. L., Ponnusamy, L., Unnasch, T. R., Hassan, H. K. & Apperson, C. S. Host-feeding patterns of Aedes albopictus (Diptera: Culicidae) in relation to availability of human and domestic animals in suburban landscapes of central North Carolina. *J. Med. Entomol.* **43**, 543–551 (2006).

59. Tchouassi, D. P. *et al.* Next generation sequencing improves the resolution of detecting mixed host blood meal sources in field collected arboviral mosquito vectors. *Med. Vet. Entomol.* **38**, 407–415 (2024).

60. Almeida, P. R. de, Weber, M. N., Sonne, L. & Spilki, F. R. Aedes-borne orthoflavivirus infections in neotropical primates - Ecology, susceptibility, and pathogenesis. *Exp. Biol. Med.* **248**, 2030–2038 (2023).

61. Sauvage, V. *et al.* New insights into Human Pegivirus-1 (HPgV-1) genotypes diversity in sub-Saharan Africa. *Infect. Genet. Evol.* **94**, 104995 (2021).

62. Shankar, E. M. *et al.* GB virus infection: a silent anti-HIV panacea within? *Trans. R. Soc. Trop. Med. Hyg.* **102**, 1176–1180 (2008).

63. Cordey, S. *et al.* Detection of dicistroviruses RNA in blood of febrile Tanzanian children. *Emerg. Microbes Infect.* **8**, 613–623 (2019).

64. Oguzie, J. U. *et al.* Metagenomic surveillance uncovers diverse and novel viral taxa in febrile patients from Nigeria. *Nat. Commun.* **14**, 4693 (2023).

65. Dastjerdi, A., Everest, D. J., Davies, H., Denk, D. & Zell, R. A novel dicistrovirus in a captive red squirrel (Sciurus vulgaris). *J. Gen. Virol.* **102**, (2021).

66. Duraisamy, R. *et al.* Detection of novel RNA viruses from free-living gorillas, Republic of the Congo: genetic diversity of picobirnaviruses. *Virus Genes* **54**, 256–271 (2018).

67. Fumwakwau, J. K., Derra, K., Nzolo, D. B., Mampunza, S. M. M. & Phanzu, D. M. Cohort profile: Kimpese health and demographic surveillance system, democratic republic of Congo. *Int. J. Epidemiol.* **53**, dyae150 (2024).

68. Sun, X. *et al.* Gut symbiont-derived sphingosine modulates vector competence in Aedes mosquitoes. *Nat. Commun.* **15**, 8221 (2024).

69. Hameed, M. *et al.* A viral metagenomic analysis reveals rich viral abundance and

520 diversity in mosquitoes from pig farms. *Transbound. Emerg. Dis.* **67**, 328–343 (2019).

521 70. Atoni, E. *et al.* Metagenomic Virome Analysis of Culex Mosquitoes from Kenya and
522 China. *Viruses* **10**, (2018).

523 71. World Health Organization. WHO position paper on dengue vaccines, May 2024. WHO:
524 Geneva, 2024. Available at:
525 https://www.who.int/publications/i/item/who-wer-9918-203-224 [Accessed 18 December
526 2024].

527

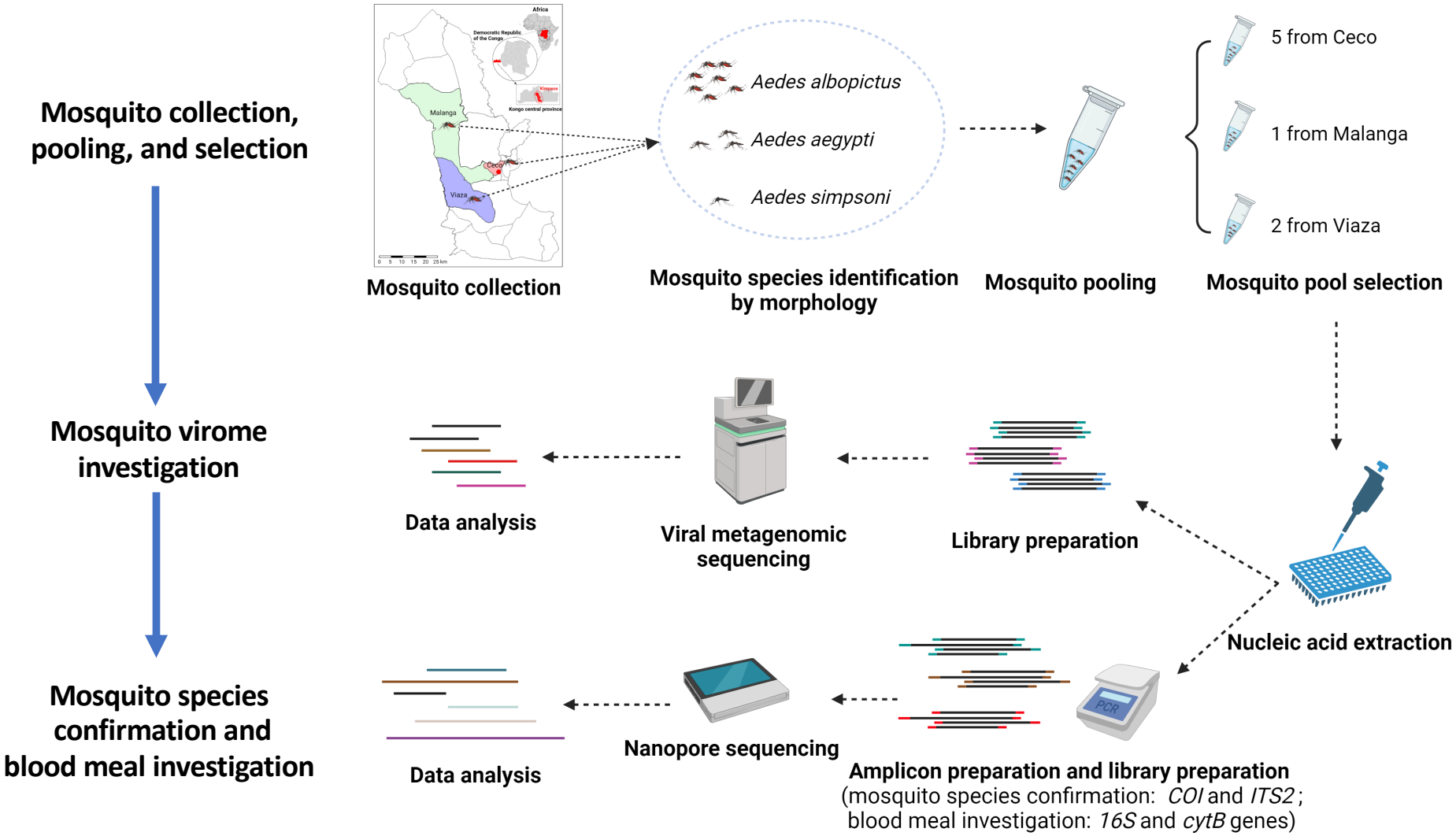
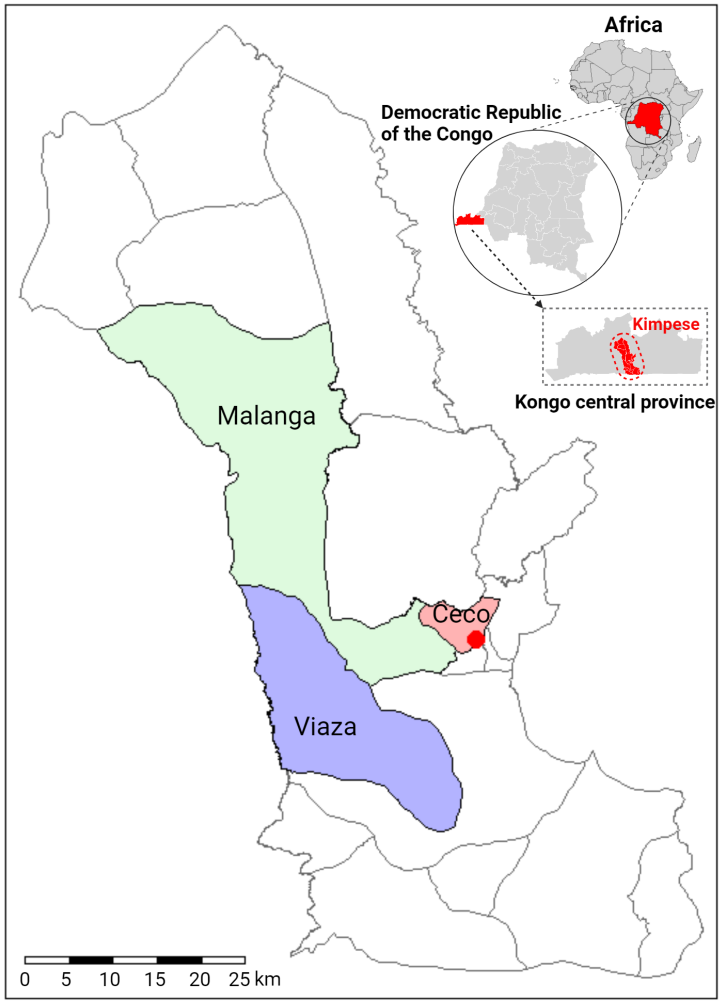



**Figure 1.** Schematic of the viral metagenomic, mosquito species, and blood meal analysis of *Aedes* mosquitoes from the Democratic Republic of the Congo.

(A)




(B)

| Sample collection site | Sample Name | Number of mosquitoes |
|---|---|---|
| Malanga | M_1 | 17 |
| Viaza | V_2 | 13 |
| Viaza | V_3 | 21 |
| Ceco | K_1 | 34 |
| Ceco | K_2 | 11 |
| Ceco | K_3 | 9 |
| Ceco | K_4 | 18 |
| Ceco | K_5 | 32 |
| Laboratory | Lab 1 | 20 |
| $H_2O$ | $H_2O$ | NA |

**Figure 2. *Aedes* mosquitoes trapped in the DRC.** A) Sample collection areas in Kimpese, DRC. Kimpese city is annotated by a red dot. B) Detailed information of the selected mosquito pools, including health areas where they were collected, and including laboratory-reared *Ae. aegypti* (Lab 1) and water ($H_2O$) controls.

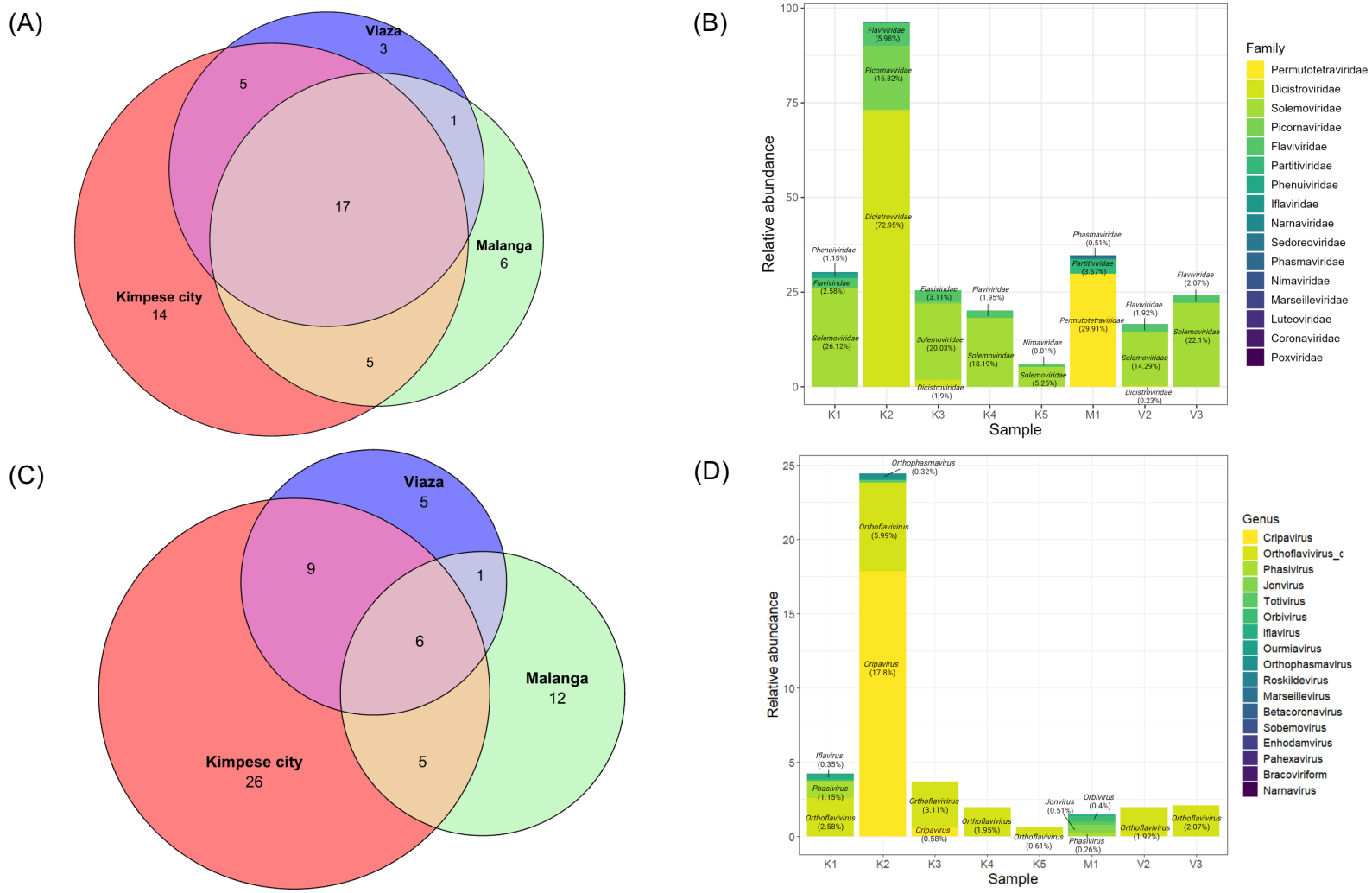



**Figure 3. Viral families and genera in field mosquito pools**. A) Venn diagram of all viral families identified in mosquito pools collected in different regions. B) Top five families in each field mosquito pool; the top three families with relative abundance >0.1% in each pool are labeled. C) Venn diagram of viral genera in mosquito pools collected in different regions (K, Kimpese city; M, Malanga; V, Viaza). B) Top five genera in field mosquito pools; the top three genus with relative abundance >0.1% in each pool are labeled. Per International Committee on Taxonomy of Viruses recommendations, relative abundance of *Orthoflavivirus* comprises both *Orthoflavivirus* and *Flavivirus* identified by Krakenuniq.

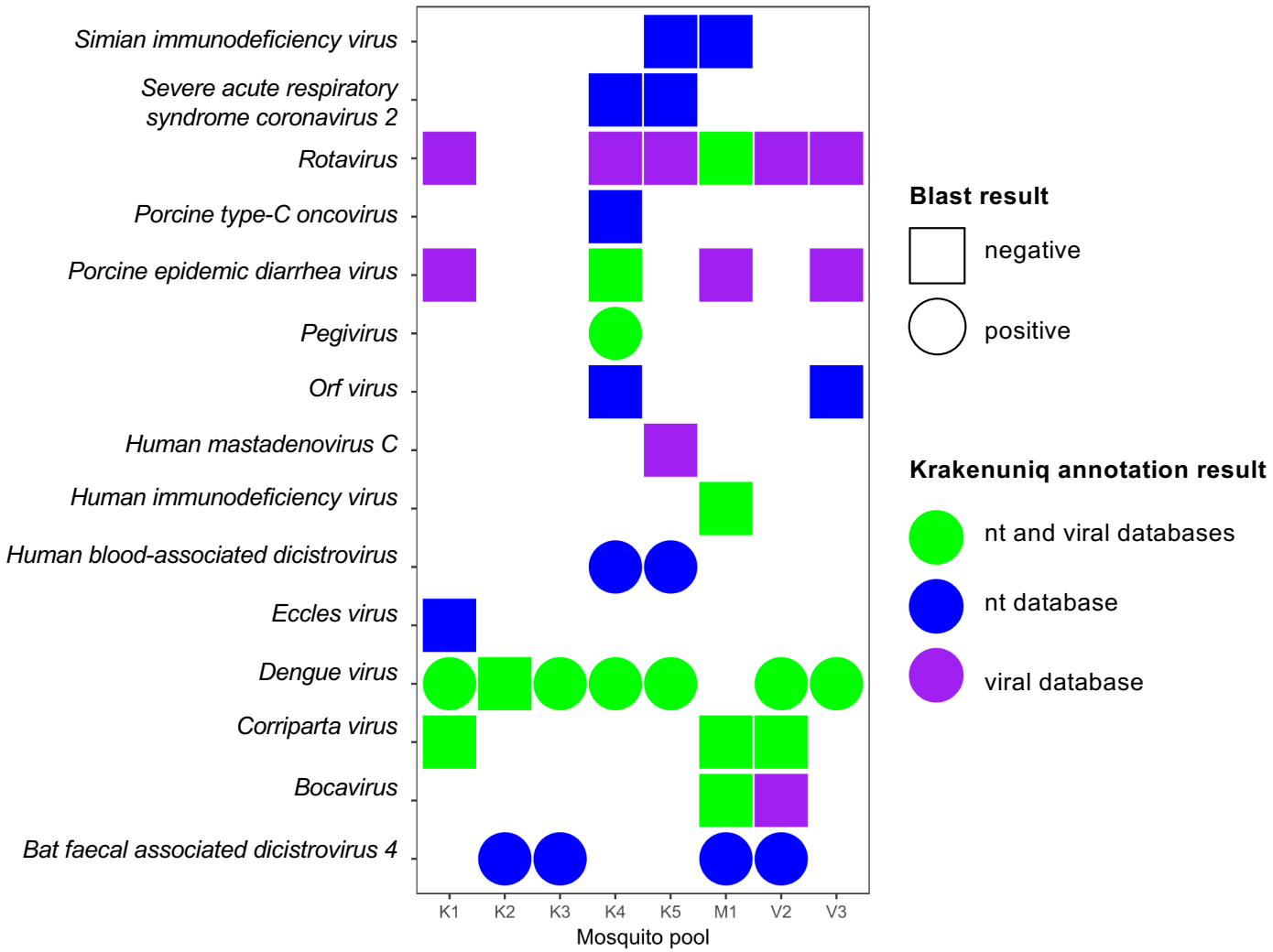



**Figure 4. Confirmation of reads mapping to known human or animal viruses.** Krakenuniq annotation results based on default nt and viral databases are shown in different colors. Blast-confirmation results for reads annotated as known human or animal viruses using krakenuniq are annotated by shape. Green circles signify confirmation by all methods used.

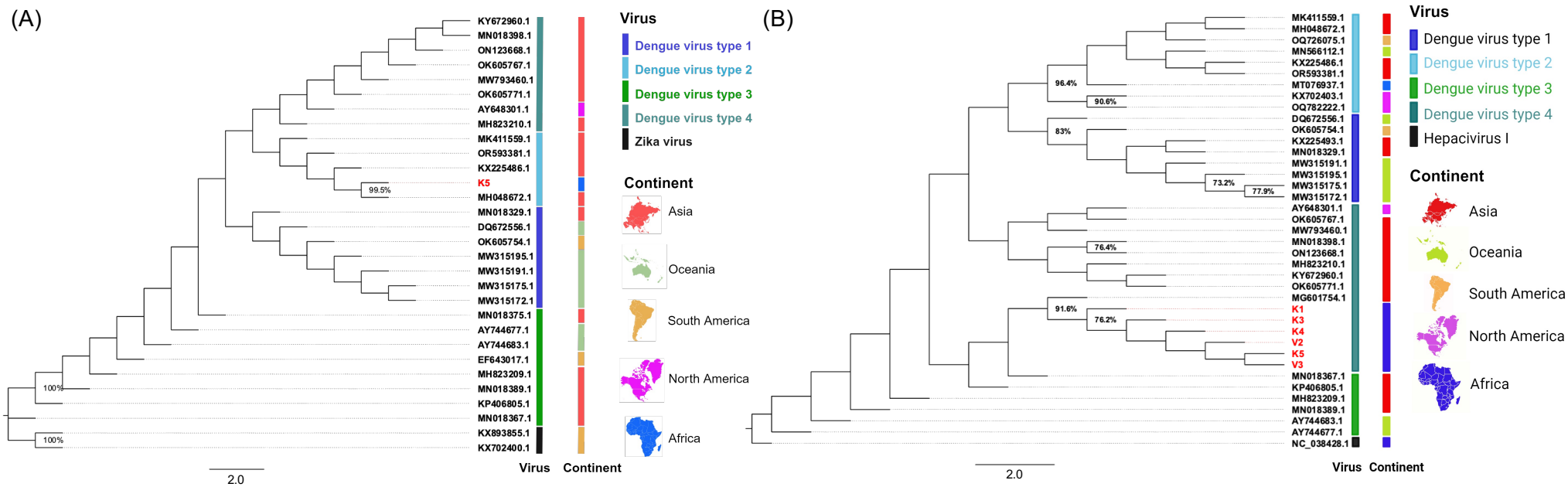



**Figure 5. Phylogenetic analysis of dengue reads from field mosquito pools.** Neighbor joining phylogenetic trees comparing DENV reads from DRC mosquito pools (red) compared to reference sequences (black). Analysis of the DENV A) non-coding region showing clustering of reads from a single DRC pool with DENV-2 reference sequences, though with limited bootstrap support, and B) 5' untranslated region (UTR) showing clustering of reads from six DRC pools with DENV-4 references sequences. Nodes with >70% bootstrap support across 1,000 replicates are annotated.

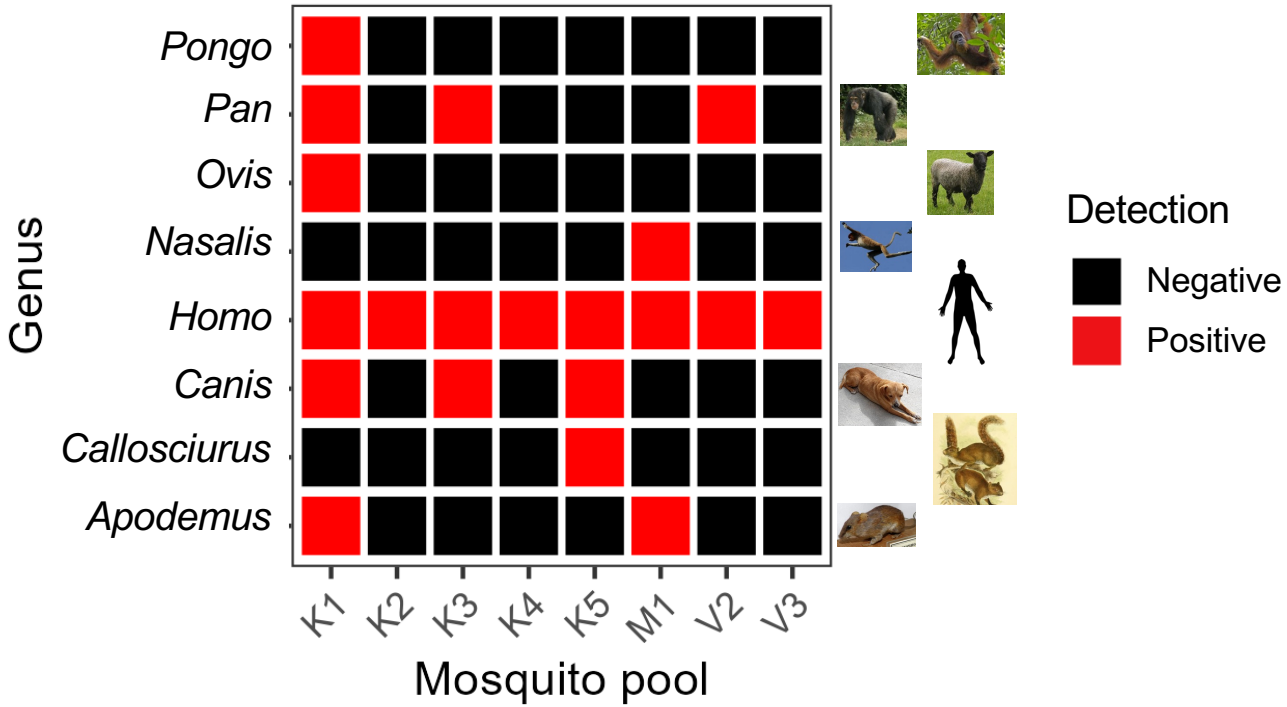



**Figure 6. Mosquito blood meal sources** determined by amplicon sequencing of the 16S ribosomal RNA and cytochrome b genes.

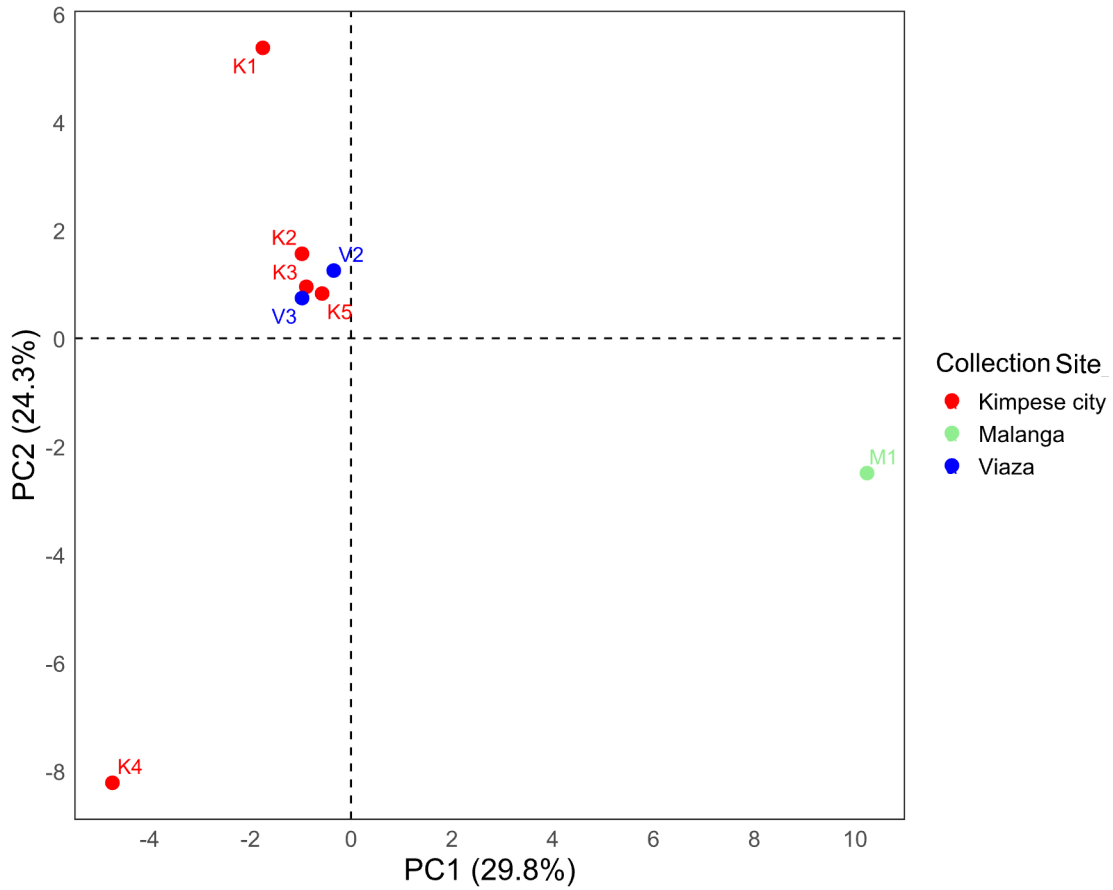



**Figure S1. Principal component analysis (PCA) annotated by collection site** using a data matrix containing relative abundance of viral genera as input.

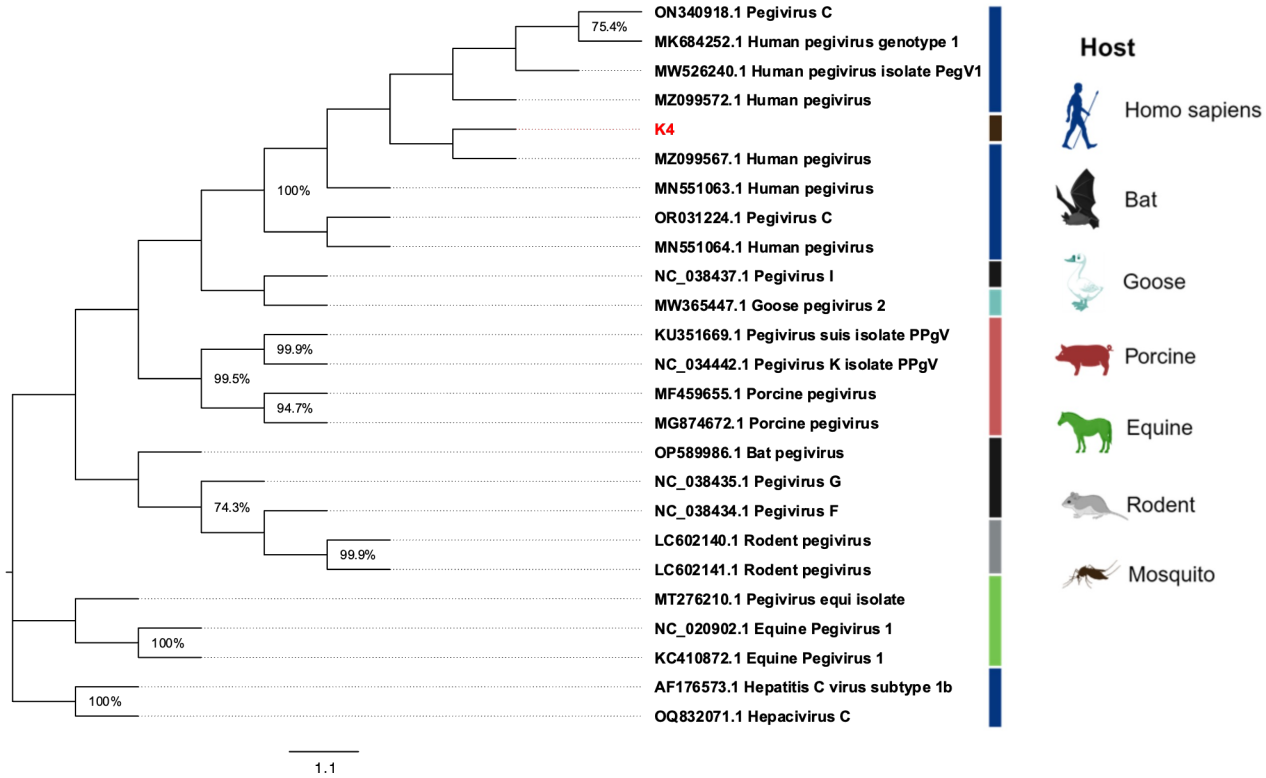



**Figure S2. Neighbor-joining phylogenetic tree of pegivirus reads** from a DRC field mosquito pool (red) compared to the polyprotein gene region of published sequences (black). Nodes with >70% bootstrap support after 1,000 replicates are annotated.

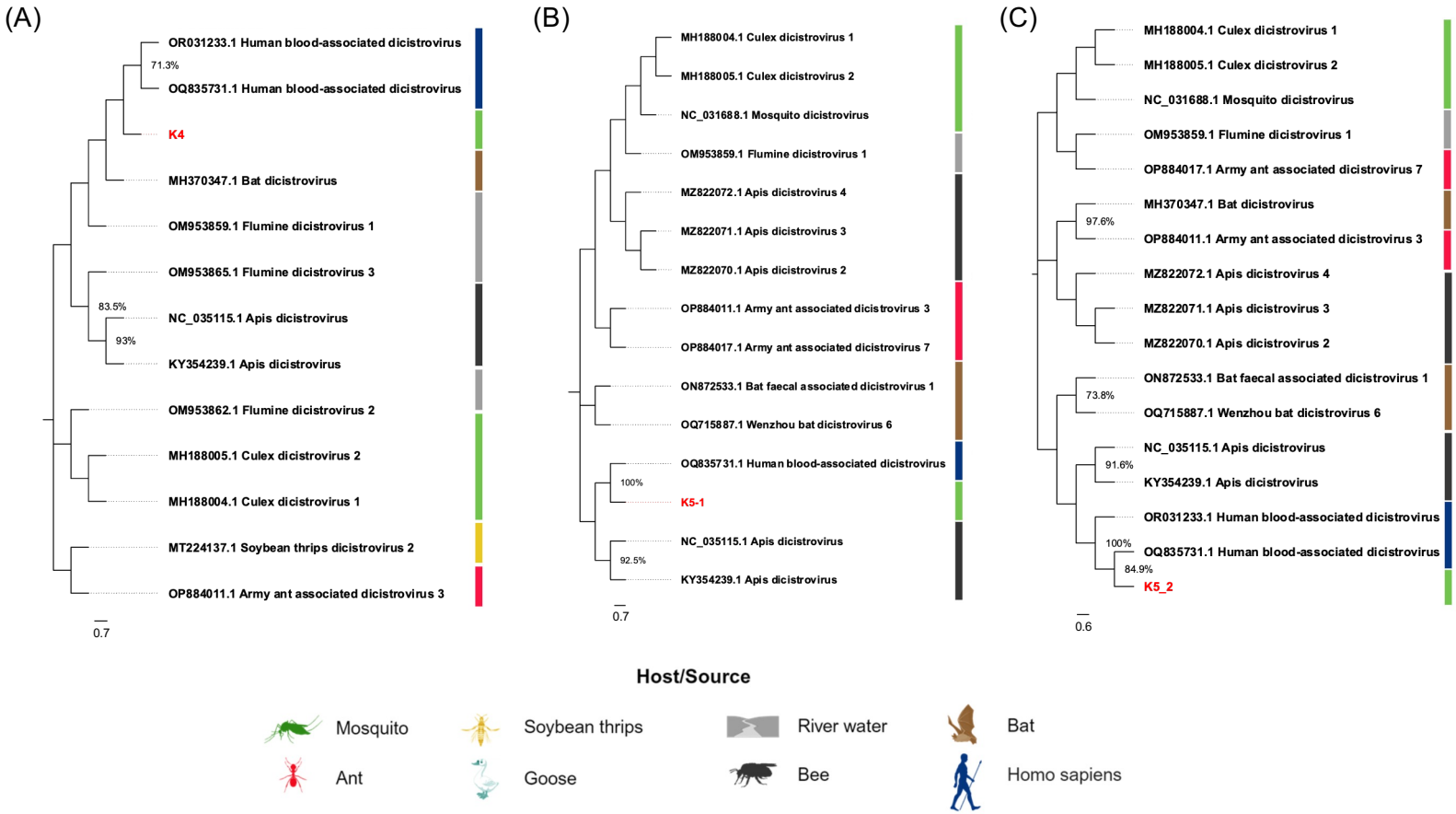



**Figure S3. Neighbor-joining phylogenetic trees of human blood-associated dicistrovirus** reads in DRC field mosquito pools (red) compared to published sequences (black). Analysis of regions of: A and B) the structural polyprotein gene and C) non-structural polyprotein gene. Nodes with >70% bootstrap support after 1,000 replicates are annotated.

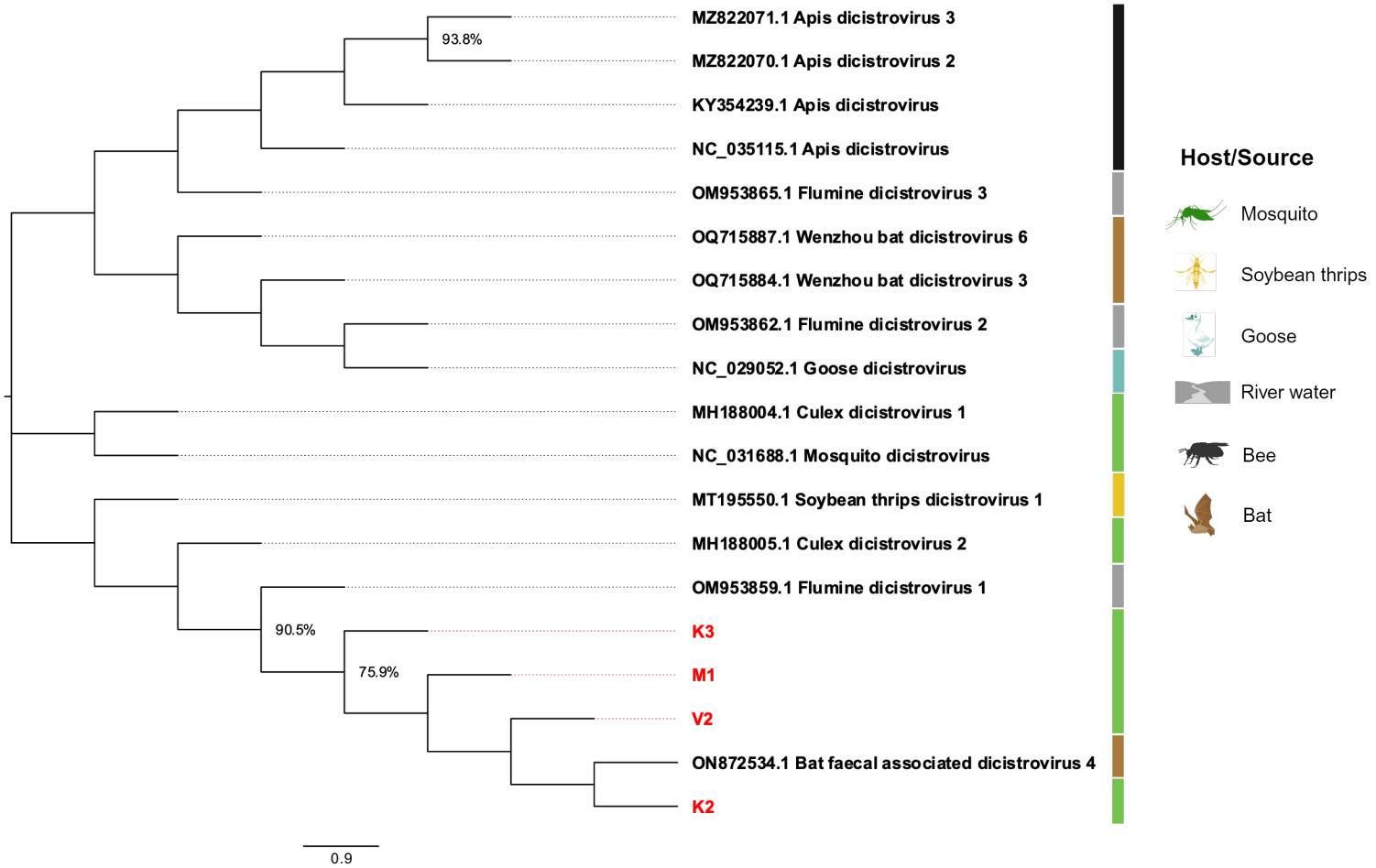



**Figure S4. Neighbor-joining phylogenetic tree of bat faecal associated dicistrovirus 4** reads in DRC field mosquito pools (red) compared to a non-coding region of published sequences (black). Nodes with >70% bootstrap support after 1,000 replicates are annotated.