Contrasted viral communities between Aedes albopictus and Culex quinquefasciatus in La Rèunion Island

Sarah François, Aymeric Antoine-Lorquin, Doriane Mutuel, Patrick Makoundou, Marco Perriat-Sanguinet, Sandra Unal, Hélène Sobry, Anne-Sophie Gosselin-Grenet, Mylène Ogliastro, Mathieu Sicard, et al.

Contrasted viral communities between *Aedes albopictus* and *Culex quinquefasciatus* in La Réunion Island

François S. [1,2], Antoine-Lorquin A. [1], Mutuel D. [1], Makoundou P. [3], Perriat-Sanguinet M. [3], Unal S. [3], Sobry H. [1], Gosselin-Grenet A.S. [1], Ogliastro M. [1], Sicard M. [3], Weill M. [3], Atyame C. [4] & Boëte C.[3*]

[1] DGIMI, Univ Montpellier, INRAE, Montpellier, France
[2] Department of Biology, University of Oxford, South Park Road, Oxford, OX1 3SY, UK
[3] ISEM, Univ Montpellier, CNRS, IRD, EPHE, Montpellier 34095, France
[4] Université de La Réunion, UMR PIMIT (Processus Infectieux en Milieu Insulaire Tropical) CNRS 9192, INSERM 1187, IRD 249, Saint-Denis, île de La Réunion, France

[*]: corresponding author: christophe.boete@umontpellier.fr

Abstract

Mosquitoes are major vectors of arboviruses, and host a wide diversity of insect viruses. Recent studies highlighted the impacts of mosquito-associated insect-specific viruses on the transmission of arboviruses. However, we still lack knowledge on the biotic and abiotic factors impacting the distribution dynamics of mosquito specific viruses, although such information has the potential to inform arbovirus surveillance efforts. To gain knowledge on the distribution of mosquito viruses in islands, we collected 13 *Aedes albopictus* and 9 *Culex quinquefasciatus* pooled larvae samples in the Réunion Island, described their whole viral communities at the family level by a viromic approach, and tested the impacts of mosquito species and spatial distance on the structure of their viral communities.
Our study show that the composition of viral communities is more strongly linked to mosquito species than to the geographic origin of samples. Spatial disparities were only observed in *Aedes albopictus* viromes. Finally, we described the genomes of five virus taxa (an iflavirus, an ambidensovirus and three potentially novel microvirus species). Notably, we detected the presence of a mosquito-infecting ambidensovirus, named CpDV, in *Aedes albopictus*. It was previously only reported in *Culex pipiens*, implying that this densovirus may have a broader host range than currently estimated.
Overall these results bring insights into the diversity and the distribution of mosquito viruses; their unexplored interactions with two major vectors of arboviral diseases warrant further studies.

Keywords:

Mosquito, viral metagenomics, virus diversity, virus surveillance, virus discovery

1. Introduction

Given their increasing global importance in public health issues, the understanding of viruses infecting mosquitoes is largely biased toward the ones that have a medical or veterinary importance. However, this has started to change with the progress of metagenomics that permits to better apprehend the diversity of insect viromes (i.e. viral communities inferred by viral genomic sequences). Apart from the fundamental interest to better understand the host-parasite interactions between mosquitoes and the viruses infecting them, such approach are considered of interest for zoonotic disease control (Olmo et *al.*, 2023). Studies of the mosquito virome allow to better estimate the diversity and the role of viruses in mosquito population dynamics, their impact on the mosquito immune system and overall how their interactions with arboviruses could eventually affect the epidemiology of the latter ones (Altinli et al., 2021). Among the newly discovered viruses, a significant proportion are considered as insect-specific viruses (ISV) and they are suspected to play a role in the transmission of arboviruses in the case of coinfection (Wei et *al*., 2006; Bolling et *al*., 2012; Hobson-Peters et *al*., 2013; Goenaga et *al*., 2015; Brinkmann, Nitsche and Kohl, 2016; Hall-Mendelin et *al*., 2016). This potential role in modulating the transmission of pathogenic viruses highlights perspectives for their eventual use as tools against arboviruses in the future (Agboli et *al*., 2019).

Despite the interest of discovering novel viruses from mosquitoes, metagenomic studies are often restricted to making inventories of the discovered viruses and, apart from a recent work (Pan et *al*., 2024), the majority of them do not address whether host taxonomy or spatial distance may constrain the geographical distribution of these newly discovered mosquito viruses, although such information has the potential to inform arbovirus surveillance efforts. Our study has attempted to overcome this limitation. Our objective was to gain knowledge on the factors impacting the circulation of mosquito viruses on the scale of an island. We tested the relative impacts of mosquito species and spatial distance on the structure of mosquito-associated viral communities. We focussed on two mosquito species that live in sympatry in Réunion Island (Boussès et *al*., 2013): *Culex quinquefasciatus* and *Aedes albopictus*.

Culex quinquefasciatus, a member of the *Culex pipiens* species complex, is a vector of parasites such as *Wuchereria bancrofti* and the avian malaria parasite *Plasmodium relictum*, but also of arboviruses such as Rift Valley fever virus, West Nile virus and Saint-Louis encephalitis virus. *C. quinquefasciatus* is distributed in tropical and subtropical regions worldwide (Samy et *al*., 2016). *Aedes albopictus* is an invasive species native from South East Asia, distributed in the intertropical areas and has spread in worldwide temperate regions (Kraemer et *al*., 2015) and it is a vector of several arboviruses such as dengue, Zika or chikungunya.

We obtained viral communities associated with 22 samples of *A. albopictus* and *C. quinquefasciatus* mosquitoes larvae that we sampled respectively in 13 and 9 locations on both the windward (east) and leeward (west) sides of Réunion Island. This island, located in the Indian Ocean, is characterised by a humid tropical climate and has been hit by a serious outbreak of Chikungunya in 2005/2006 (Josseran et *al*., 2006). Chikungunya disease is still present in Réunion Island in 2024, where dengue is also becoming endemic (Hafsia et *al*., 2022; Vincent et *al*., 2023).

2. Material and Methods

2.1. Sampling of Aedes albopictus and Culex *quinquefasciatus populations*

Larvae of *A. albopictus* and *C. quinquefasciatus* were collected between February and April 2019 in 14 larvae breeding sites across Réunion Island (Fig. 1). Morphological characteristics (body shape, siphon, antennae) were used for species identification. Almost all sites (except the Sainte-Marie Ravine site) were cemeteries, corresponding to artificial breeding sites, where mosquito larvae grow in flower pots that are regularly supplied with water (Table 1). For each site and for each species, a batch of larvae was collected, rinsed several times with tap water and larvae were sorted according to their species. Samples were then kept in Eppendorf tubes and transported on ice to the lab then stored at -80°C until processing. Each sample was a pool of larvae (stage L3/L4) and weighted between 1.17g and 1.49g for *A. albopictus* and between 0.5g and 1.42g for *C. quinquefasciatus* (Table 1). Both mosquito species were in sympatry in 8 sites, *A. albopictus* was found alone in 5 sites and *C. quinquefasciatus* was alone in one site (Fig. 1).

2.2. Sample preparation and sequencing

Viromes were obtained from 22 mosquito samples (Table 1) as described in (François et *al*., 2018). Briefly, 1.35 g ±0.1 of material per sample of *A. albopictus* and 1.1 g ±0.3 of material per sample for *C. quinquefasciatus* was processed using a virion-associated nucleic acid (VANA) based metagenomics approach to screen for the presence of both DNA and RNA viruses. Samples were grounded and centrifuged to recover supernatants that were filtered through a 0.45 µm filter and centrifuged at 140,000 g for 2.5 hours to concentrate viral particles. The resulting pellets were resuspended, and nucleic acids not protected in virus-like particles (VLPs) were degraded by DNase and RNase incubation at 37°C for 1.5 h. Total RNA and DNA were then extracted using a NucleoSpin kit (Macherey Nagel, Bethlehem, PA, USA) following manufacturer's instructions. Reverse transcription was performed by SuperScript III reverse transcriptase (Invitrogen), cDNAs were purified by a QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and complementary strands synthesised by Klenow DNA polymerase I. Double-stranded DNA was amplified by random PCR amplification. Samples were barcoded during reverse transcription and PCR steps using homemade 26-nt Dodeca Linkers coupled to homemade complementary PCR multiplex identifier primers. PCR products were purified using NucleoSpin gel and PCR clean-up kit (Macherey Nagel, Bethlehem, PA, USA) following manufacturer's instructions. Finally, libraries were prepared using NEBNext Ultra DNA PCR free with Ilumina adapter kit without fragmentation step from purified amplicons and sequenced on an Illumina HiSeq 3000 to generate 2x150bp paired-end reads (Genewiz, South Plainfield, NJ, USA).

2.3. Viral Sequence analysis and Genome reconstruction

Illumina adaptors were removed and reads were filtered for quality (q30 quality and read length >45 nt) using cutadapt 2.19 (Martin, 2011). Cleaned reads were assembled *de novo* into contigs using MEGAHIT 1.2.9 (Li et al., 2015). Taxonomic assignment was achieved on contigs of length >900 nt through searches against the NCBI gbvrl viral database (created on 07/07/2021) using DIAMOND 0.9.30 with an e-value cutoff of $<10^{-3}$ (Buchfink et al., 2015). All contigs that matched virus sequences were selected and used as queries to perform reciprocal searches on NCBI non-redundant protein sequence database (created on July 2020) with an e-value cutoff of $<10^{-3}$ in order to eliminate false positives.

Viral contigs completion and coverage was assessed by iterative mapping using Bowtie2 3.5.1 with the options end−to-end and very-sensitive (Langmead, 2010). Putative Open Reading Frames (ORFs) were identified using ORF finder (length cutoff >400 nt) on Geneious prime 2021.1.1 (Kearse et *al*., 2012). Microviruses genome circularisation was performed using Simple-Circularise 1.0 script (https://github.com/Kzra/Simple-Circularise) with minimal overlap length of 20nt, and genome coverage was assessed by iterative mapping using Bowtie2 3.5.1 with end-to-end sensitive options (Langmead, 2010). ORFs were annotated by blastn query-centred alignment against the complete NCBI RefSeq viral database (created on 23/04/2021). Viral genomes completion was verified manually by aligning them to their 10 closest relatives downloaded from the GenBank nucleotide database (nr). In all subsequent analyses, we focused only on full coding sequences (100% of CDS) based on the alignments of genomes with their closest relatives combined with ORF completeness, thus discarding contigs with partial CDS. The viral isolates that belonged to already described species were reconstructed as follows: after mapping against the closest relative deposited in the NCBI nucleotide database, consensus sequences were generated using samtools 1.2 (Danecek et *al*., 2021). Mutations were called at depth ≥5 if they differed from the reference isolate; otherwise, sites were kept as those of the reference isolate.

2.4. Viral Discovery and Taxonomic Assignment

To determine if viral contigs belonged to new species, their nucleotide sequences or their predicted protein sequences were aligned and compared with the 10 closest related viral genomes found by similarity searches performed above using MAFFT 7.450 with the G-INS-i algorithm (Katoh et *al*., 2002; Katoh & Standley, 2013) or MUSCLE 3.8.425 (16 iterations) (Edgar, 2004) using default settings. Genomes were classified as new virus strain or new virus species according to the species demarcation thresholds recommended within the online reports of the International Committee on Taxonomy of Virus (ICTV, https://ictv.global/).

2.5. Phylogenetic Analyses

Phylogenetic trees were built using Maximum Likelihood methods for all reconstructed full-length CDS genomes in order to place them within the currently known viral diversity and to infer their possible host range. Representative sets of replication-associated and capsid proteins, or polyproteins, were extracted from the NCBI GenBank non-redundant (nr) database for each taxonomic group in which the genomes were classified, comprising all the ICTV ratified species and the 10 closest relatives found by blastx search against GenBank nr DB on 27/06/2024, expect for the *Densovirinae* phylogenetic tree that was based on all the *Protoambidensovirus dipteran1* lineages.
Viral nucleotide sequences were aligned using MAAFT (option G-INS-i). Phylogenetic trees were constructed in RAxML 8.2.11 (Stamatakis, 2014) using the GTR +GAMMA + I nucleotide evolution model. Amino acid sequences were aligned using MUSCLE 5.1 (PPP algorithm) with default settings. Sequences that were not reliably aligned due to high amino acid divergence were removed and the dataset subsequently realigned. Phylogenetic trees were constructed in RAxML 8.2.11 (Stamatakis, 2014) using the LG+I+G protein evolution model. Tree branch support was estimated using 100 bootstrapped replicates. The trees were mid-point rooted and visualised with MEGAX 10.2.4 (Stecher et *al*., 2020).

2.6. Statistical Analyses

Statistical analyses were performed using R and RStudio version 1.1.456 softwares (R Core Team, 2023). Data from contingency tables were standardised to allow inter-sample comparisons: taxonomic binning artefacts and potential inter-sample contamination were restricted by applying an abundance threshold >1/100,000 reads/taxon/sample, and, for taxa that contaminated negative controls, by removing them from sample datasets where their abundance was equal or inferior to their abundance in controls (i.e. the number of reads associated to a given family had to be above a threshold of 800 reads corresponding to the background noise that was detected in the negative control).
To enable comparison between viral taxa that have different genome lengths, the number of virus reads was divided by the length of the viral contig to which it mapped (kb). Viral diversity accumulation curves in mosquito viromes were made using the Vegan package (*vegan*, 2012). All the following analyses on virus community richness and composition were conducted at the viral family level. The impact of host species on viral community α-diversity was evaluated using the Shannon and Simpson diversity indices, and on β-diversity using Bray-Curtis dissimilarity index. Differences in viromes structure of *A. albopictus* and *C. quinquefasciatus* were visualised using a heatmap ("ggplot2: Elegant Graphics for Data Analysis (3e)"). The effect of host species on viral communities was also determined by one-factor PERMANOVAs with 10,000 permutations on Bray–Curtis matrix, using the adonis function of the VEGAN package (Dixon, 2003). Permutational tests of dispersions (PERDISPs) using the function permutest.betadisper (10000 permutations, pairwise) were performed to assess whether significant effects could be influenced by differences in group dispersion (Anderson, 2001). Statistical significance of PERMANOVA results was assumed when $p < 0.01$. The impact of host species on viral abundance was assessed using Wilcoxon tests. Rare taxonomic groups (occurring in <5 samples) were not considered for abundance analysis. Multiple comparison test adjustment of p values was performed using the Benjamini-Yekutieli (BY) method (Benjamini & Yekutieli, 2001). Finally, a potential correlation between sampling sites spatial distances and virome composition was assessed by a Mantel test (10,000 random permutations and Pearson correlation, package ade4 (Dray & Dufour, 2007)) on Bray-Curtis virome beta diversity dissimilarity matrix.

3. Results

3.1. Viromes comparison between sympatric *Culex quinquefasciatus* and *Aedes albopictus*, and virus spatial distribution in the Réunion island

The sequencing depth was sufficient to recover the entire viral communities at the family level as indicated by the rarefaction curves (Fig. 2a). In addition, the accumulation curve (number of taxa/number of samples) (Fig. 2b) revealed that our sampling effort permitted to recover the entire diversity of viral families circulating in the targeted Culicidae samples.

We reconstructed 46 viral contigs (ST1) ranging from 314 to 10844 nt. The taxonomic assignment of the contigs obtained from the 22 samples (13 *A. albopictus* samples and 9 *C. quinquefasciatus* samples) to viral families was highly variables among samples, ranging between 0 to 100 % with an average value of 59.61% of the total number of cleaned reads assigned to viral contigs, the remaining reads not being classified (Fig. 3). This range of viral

read abundance is congruent with another recent insect viral metagenomic studies based on the same protocol (François et *al*., 2021). We detected viral sequences that could not be classified at the family level, as they belonged to unclassified virus clades. These unclassified viral sequences were detected in all samples, with the exception of *C. quinquefasciatus* from Saint-Benoît, and their abundance was higher (>90% reads) in *A. albopictus* from La Possession and *C. quinquefasciatus* from Saint-Denis. As they could not be classified into existing families, these sequences were not taken into account in further analyses.

Overall, we found 12 virus families (including 6 associated with *A. albopictus* and 9 with *C. quinquefasciatus*) in the larvae pools (Fig. 4). Notably, a majority of families (8/12) have RNA genomes. Among them, 2 families of positive single-stranded RNA viruses were known as insect-specific viruses (*Dicistroviridae* and *Iflaviridae*); while only one of DNA viruses (*Parvoviridae*) was known to infect insects (*Densovirinae* subfamily). Arthropod-infecting viruses represented 67.9% of all reads, and the most represented families in term of read abundance are *Parvoviridae* (single-stranded DNA virus, 44.7% of all virus reads) and *Iflaviridae* (positive single-stranded RNA virus, 21.1% reads). Members of the *Phasmaviridae* family (negative single-stranded RNA virus whose host range includes insects) are less abundant (around 1% of the total number of reads in *Culex* samples), and were detected in 4 sites. Sequences related to bacteriophages families were also detected (*Microviridae*, ST2) and account for 4.2% of the total viral reads.

Culex samples (mean number of viral reads: 96,296; min: 619 reads; max: 532,533 reads, average percentage of viral reads: 4.2%) did not contain more virus reads than *Aedes* samples (mean number of viral reads: 25,350; min: 727; max: 225,372, average percentage of viral reads: 15.4%), Wilcoxon test *p* value = 0.14. Considering the most represented families in term of occurrence, 92.1% of the reads are related to 5 families; and viruses of the family *Parvoviridae* were found in 3 *Aedes* samples and in 2 *Culex* samples, while iflavirus reads were detected in 5 *Aedes* samples and in 7 *Culex* samples (Fig. 4, ST2).

We then evaluated whether mosquito species could explain for differences in mosquito viromes richness and composition. While we found no differences in Shannon and Simpson α-diversity indices between mosquito species (SF1), MDS plot and PERMANOVAs based on Bray-Curtis dissimilarity matrices showed significant differences in virus community composition between mosquito species (*p* value = 0.0001 and R2 = 17 %, Fig. 5). We conducted a differential abundance analysis by Wilcoxon tests corrected by BY method for multiple comparisons on viral families present in > 5 samples (*i.e. Iflaviridae, Microviridae, Parvoviridae, Phasmaviridae, Phycoviridae, Rhabdoviridae* and *Siphoviridae* families) to assess which viruses drove this observed difference (Fig. 6). Difference in community composition was driven by *Microviridae* family members (bacteriophages) that were more abundant in *A. albopictus* samples, and by insect-infecting viruses belonging to the *Phasmaviridae* and *Rhabdoviridae* families which were more abundant in *C. quinquefasciatus* samples (adjusted *p* value < 0.05).
Finally, we tested whether spatial distance between sampling locations could impact mosquito viromes composition, by conducting a Mantel test using Pearson correlation based on Bray-Curtis dissimilarity matrix and a pairwise distance matrix between sampling locations.

While Mantel test showed no significance for *C. quinquefasciatus* viromes, it showed a negative correlation between geographical distance and viromes similarity for *A. albopictus* samples (*p* value = 0.016, observed correlation: 31%).
Our study thus suggests that there were differences of virus distribution between the two mosquito species, and spatial variation in *A. albopictus* viromes composition. Our results tend to show that virome composition may be more explained by mosquito species than by spatial distribution that may be linked to unmeasured ecological conditions (Fig. 5 and 6).

3.2. Virus discovery and phylogenetic analysis

Incomplete viral coding sequences (CDS) were discarded from the following analyses, as only virus with complete CDS can be classified as belonging to new or already known species using ICTV species demarcation criteria based on sequence data similarity to their closest relatives (Simmonds et *al.*, 2017; Siddell et *al.*, 2023).
We reconstructed 5 viral full-length CDS or genomes. The coverage depth across all these genomes is high (between 62 and 12,521) with read numbers ranging from 2,200 to 501,573 mapped reads per sample for each genome (Table 2). We built ML phylogenetic trees to place these viruses within the currently known viral diversity and to infer their possible host range.

3.2.1. Densovirinae

Densovirinae is a subfamily of the *Parvoviridae* family whose members infect arthropods, crustaceans and echinoderms (Cotmore et *al.*, 2014). We report the complete CDS of a densovirus, named Culex quinquefasciatus associated densovirus isolate 2019_VP12-D87, whose contig length is 5,459nt, and which shows 89.0% genome-wide nucleotide identity to a *Dipteran protoambidensovirus 1* lineage isolated from *Culex pipiens* mosquitoes (FJ810126.1, *Protoambidensovirus* genus, common virus name: Culex pipiens densovirus) (Table 2). According to the species demarcation threshold in the *Densovirinae* subfamily (i.e., < 85% NS1 protein identity (Cotmore et *al.*, 2014)), this new genome (accession number: PQ041300) thus represents a new distant lineage of *Dipteran protoambidensovirus*. The position of our CpDV sequence in the NS1 gene tree shows that our lineage belongs to the CpDV-3 clade which was previously only represented by sequences collected from Beijing (China) samples in 2003 (Fig. 7) (Altinli et *al.*, 2019).

3.2.2. Iflaviridae

The *Iflaviridae* family comprises picorna-like viruses infecting arthropods. In our study we found a novel iflavirus linage which was detected in a pool of *C. quinquefasciatus* from Saint-Benoît. We named it Culex quinquefasciatus associated iflavirus isolate Saint_Denis, whose contig length is 10,074nt (Table 2, accession number: PQ041301). Its closed relative is the *XiangYun picorna-like virus 4* species (Feng et al., 2022) that was discovered in *Culex theileri* from Yunnan, China (Iflavirus genus, accession number OL700176) with which it shared a whole polyprotein pairwise identity of 98.5 (Table 2). The *Iflaviridae* species demarcation criterion is set as < 90% of capsid proteins identity (Valles et *al.*, 2017). Our virus thus belongs to the *XiangYun picorna-like virus 4* species, which may infect *Culex spp*. Finally, Culex quinquefasciatus associated iflavirus isolate 2019_VP12-D85 clustered in a monophyletic

clade containing only viral taxa isolated from dipteran species (mosquitoes and true flies), indicating that this lineage is likely specific of dipterans (Fig. 8).

3.2.3. Microviridae

We reconstructed 3 complete circularised genomes belonging to bacteriophages from the *Microviridae* family (circular ssDNA viruses). Their similarity to already known microvirus species is relatively low (54.5% to 75.0% of amino acid identity in major capsid and replication-associated proteins) (Table 2). Thus, while there is currently no species demarcation criteria based on genome similarity defined by the ICTV for the *Microviridae* family (Kirchberger & Ochman, 2023), those 3 genomes could belong to novel viral species, tentatively named Aedes albopictus associated microvirus 1, 2 and 3 (accession numbers: PQ041302, PQ041303 and PQ041304, respectively). Those 3 microviruses cluster with members of the *Gokushovirinae* subfamily that were isolated from water ecosystems (Fig. 9, Table 2).

4. Discussion

We studied the viromes of two mosquito species that are major vectors of arboviruses. However, we did not detect any arboviruses in our screened larvae samples despite the ongoing circulation of dengue in Réunion Island; this is explained by the fact that we studied the virome isolated from larvae and not from adult mosquitoes and by the low level of vertical transmission of arboviruses, which is typically between 0.1 and 4 % (Adams & Boots, 2010; Lequime et *al.*, 2016). Our study shows that, in the Réunion Island, patterns of mosquito-associated virus distribution may be more explained by mosquito species rather than by geographical location. Our results are congruent with those obtained from other viral metagenomic studies (Sadeghi et *al.*, 2018; Sanborn et *al.*, 2019; Kubacki et *al.*, 2020; Thongsripong et al., 2021), which highlighted that virus distributions is correlated to the mosquito taxonomy (Faizah et *al.*, 2020), as shown previously in a metatranscriptomic study conducted in China (Pan et *al.*, 2024) and a shotgun sequencing study from Thailand (Thongsripong et *al.*, 2021).

We used a stringent threshold for virus detection, and thus only classified complete viral CDS at the species level, in order to reduce the risk of false-positive detection errors, at the cost of underestimating the number of virus taxa present in our viromes. However, we could reconstruct the full-length genomes of three potential new bacteriophages species belonging to the *Microviridae* family. Their specifically high abundance in *Aedes* samples suggests than they may infect *Aedes albopictus* larvae microbiota, although their presence could also be due to environmental or diet contamination by infected aquatic bacteria. However their absence or low abundance in the *Culex* samples (some of them sharing the same breeding sites with the *Aedes*) tends to favour the idea of an association with *Aedes albopictus*.
We also reported the full CDS of a divergent insect-infecting iflavirus lineage belonging to the *XiangYun picorna-like virus 4* species. This iflavirus clustered in a monophyletic clade containing only viral taxa isolated from dipteran species (mosquitoes and true flies), and was previously reported by Feng et *al.* in *Culex theileri* in China (Feng et *al*., 2022). Altogether, these results indicate that this virus is probably common in *Culex spp*. Further work is needed to determine the prevalence of XiangYun picorna-like virus 4 in natural populations of mosquitoes and its eventual impact on mosquito fitness.

Interestingly, we finally reported for the first time the presence of *Dipteran protoambidensovirus 1* (common name: Culex pipiens densovirus, clade CpDV-3), a mosquito-infecting ambidensovirus, in *A. albopictus.* Reported CpDV host range currently only includes *Culex pipiens* (Altinli et *al.*, 2019, 2020), its broadening to *A. albopictus* species warrants further confirmation, as a broader host range likely impacts CpDV evolution.

Data availability statement

The genomic sequences of the five full-length viral genomes or CDS have been deposited at GenBank under the accession numbers PQ041300 to PQ041304. High-throughput sequencing reads were deposited in SRA under the accession no. SRR29133481 to SRR29133504 under PRJNA1114772 BioProject. The bioinformatics pipeline code is available at https://github.com/ayantoine/NearVANA.

References

Adams, B., Boots, M. (2010). How important is vertical transmission in mosquitoes for the persistence of dengue? Insights from a mathematical model. *Epidemics*, 2, 1–10. https://doi.org/10.1016/j.epidem.2010.01.001

Agboli, E., Leggewie, M., Altinli, M., Schnettler, E. (2019). Mosquito-specific viruses—transmission and interaction. *Viruses*, 11, 873. https://doi.org/10.3390/v11090873

Altinli, M., Lequime, S., Atyame, C., Justy, F., Weill, M., Sicard, M. (2020). Wolbachia modulates prevalence and viral load of *Culex pipiens* densoviruses in natural populations. *Molecular Ecology*, 29, 4000–4013. https://doi.org/10.1111/mec.15609

Altinli, M., Lequime, S., Courcelle, M., François, S., Justy, F., Gosselin-Grenet, A.-S., Ogliastro, M., Weill, M., Sicard, M. (2019). Evolution and phylogeography of *Culex pipiens* densovirus. *Virus Evolution*, 5, vez053. https://doi.org/10.1093/ve/vez053

Altinli, M., Schnettler, E., Sicard, M. (2021). Symbiotic interactions between mosquitoes and mosquito viruses. *Frontiers in Cellular and Infection Microbiology*, 11. https://doi.org/10.3389/fcimb.2021.694020

Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26, 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Benjamini, Y., Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29, 1165–1188. https://doi.org/10.1214/aos/1013699998

Bolling, B.G., Olea-Popelka, F.J., Eisen, L., Moore, C.G., Blair, C.D. (2012). Transmission dynamics of an insect-specific flavivirus in a naturally infected *Culex pipiens* laboratory colony and effects of co-infection on vector competence for West Nile virus. *Virology*, 427, 90–97. https://doi.org/10.1016/j.virol.2012.02.016

Boussès, P., Dehecq, J.S., Brengues, C., Fontenille, D. (2013). Inventaire actualisé des moustiques (Diptera: Culicidae) de l'île de La Réunion, océan Indien. *Bulletin de la Société de pathologie exotique*, 106, 113–125. https://doi.org/10.1007/s13149-013-0288-7

Brinkmann, A., Nitsche, A., Kohl, C. (2016). Viral metagenomics on blood-feeding arthropods as a tool for human disease surveillance. *International Journal of Molecular Sciences*, 17, 1743. https://doi.org/10.3390/ijms17101743

Buchfink, B., Xie, C., Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12, 59–60. https://doi.org/10.1038/nmeth.3176

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Cotmore, S.F., Agbandje-McKenna, M., Chiorini, J.A., Mukha, D.V., Pintel, D.J., Qiu, J., Soderlund-Venermo, M., Tattersall, P., Tijssen, P., Gatherer, D., Davison, A.J. (2014). The family Parvoviridae. *Archives of Virology*, 159, 1239–1247. https://doi.org/10.1007/s00705-013-1914-1

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10, giab008. https://doi.org/10.1093/gigascience/giab008

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, 14, 927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x

Dray, S., Dufour, A.-B. (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software*, 22, 1–20. https://doi.org/10.18637/jss.v022.i04

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

 Faizah, A. N., Kobayashi, D., Isawa, H., Amoa-Bosompem, M., Murota, K., Higa, Y., Futami, K., Shimada, S., Kim, K. S., Itokawa, K., Watanabe, M., Tsuda, Y., Minakawa, N., Miura, K., Hirayama, K., & Sawabe, K. (2020). Deciphering the virome of *Culex vishnui* subgroup mosquitoes, the major vectors of Japanese encephalitis, in Japan. *Viruses*, 12, 264. https://doi.org/10.3390/v12030264

 Feng, Y., Gou, Q.-Y., Yang, W.-H., Wu, W.-C., Wang, J., Holmes, E. C., Liang, G., & Shi, M. (2022). A time-series meta-transcriptomic analysis reveals the seasonal, host, and gender structure of mosquito viromes. *Virus Evolution*, 8, veac006. https://doi.org/10.1093/ve/veac006

 François, S., Antoine-Lorquin, A., Kulikowski, M., Frayssinet, M., Filloux, D., Fernandez, E., Roumagnac, P., Froissart, R., & Ogliastro, M. (2021). Characterisation of the viral community associated with the alfalfa weevil (*Hypera postica*) and its host plant, alfalfa (*Medicago sativa*). *Viruses*, 13, 791. https://doi.org/10.3390/v13050791

 François, S., Filloux, D., Fernandez, E., Ogliastro, M., & Roumagnac, P. (2018). Viral metagenomics approaches for high-resolution screening of multiplexed arthropod and plant viral communities. In Pantaleo, V. & Chiumenti, M. (Eds.), *Viral Metagenomics: Methods and Protocols* (pp. 77–95). Springer, New York, NY. https://doi.org/10.1007/978-1-4939-7683-6_7

 Goenaga, S., Kenney, J. L., Duggal, N. K., Delorey, M., Ebel, G. D., Zhang, B., Levis, S. C., Enria, D. A., & Brault, A. C. (2015). Potential for co-infection of a mosquito-specific flavivirus, Nhumirim virus, to block West Nile virus transmission in mosquitoes. *Viruses*, 7, 5801–5812. https://doi.org/10.3390/v7112911

 Hafsia, S., Haramboure, M., Wilkinson, D. A., Baldet, T., Yemadje-Menudier, L., Vincent, M., Tran, A., Atyame, C., & Mavingui, P. (2022). Overview of dengue outbreaks in the southwestern Indian Ocean and analysis of factors involved in the shift toward endemicity in Reunion Island: A systematic review. *PLOS Neglected Tropical Diseases*, 16, e0010547. https://doi.org/10.1371/journal.pntd.0010547

 Hall, R. A., Bielefeldt-Ohmann, H., McLean, B. J., O'Brien, C. A., Colmant, A. M. G., Piyasena, T. B. H., Harrison, J. J., Newton, N. D., Barnard, R. T., Prow, N. A., Deerain, J. M., Mah, M. G. K. Y., & Hobson-Peters, J. (2016). Commensal viruses of mosquitoes: Host restriction,

transmission, and interaction with arboviral pathogens. *Evolutionary Bioinformatics*, 12s2, EBO.S40740. https://doi.org/10.4137/EBO.S40740

Hobson-Peters, J., Yam, A. W. Y., Lu, J. W. F., Setoh, Y. X., May, F. J., Kurucz, N., Walsh, S., Prow, N. A., Davis, S. S., Weir, R., Melville, L., Hunt, N., Webb, R. I., Blitvich, B. J., Whelan, P., & Hall, R. A. (2013). A new insect-specific flavivirus from Northern Australia suppresses replication of West Nile virus and Murray Valley encephalitis virus in co-infected mosquito cells. *PLOS ONE*, 8, e56534. https://doi.org/10.1371/journal.pone.0056534

Josseran, L., Paquet, C., Zehgnoun, A., Caillere, N., Le Tertre, A., Solet, J.-L., & Ledrans, M. (2006). Chikungunya disease outbreak, Reunion Island. *Emerging Infectious Diseases*, 12, 1994–1995. https://doi.org/10.3201/eid1212.060710

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30, 3059–3066. https://doi.org/10.1093/nar/gkf436

Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30, 772–780. https://doi.org/10.1093/molbev/mst010

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kirchberger PC, Ochman H (2023) Microviruses: A World Beyond phiX174. *Annual Review of Virology* 10, 99–118. https://doi.org/10.1146/annurev-virology-100120-011239

Kraemer MUG, Sinka ME, Duda KA, Mylne AQN, Shearer FM, Barker CM, Moore CG, Carvalho RG, Coelho GE, Van Bortel W, Hendrickx G, Schaffner F, Elyazar IRF, Teng H-J, Brady OJ, Messina JP, Pigott DM, Scott TW, Smith DL, Wint GRW, Golding N, Hay SI (2015) The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *eLife* 4, e08347. https://doi.org/10.7554/eLife.08347

Kubacki J, Flacio E, Qi W, Guidi V, Tonolla M, Fraefel C (2020) Viral Metagenomic Analysis of *Aedes albopictus* Mosquitos from Southern Switzerland. *Viruses* 12, 929. https://doi.org/10.3390/v12090929

Langmead B (2010) Aligning Short Sequencing Reads with Bowtie. *Current Protocols in Bioinformatics* 32, 11.7.1-11.7.14. https://doi.org/10.1002/0471250953.bi1107s32

Lequime S, Paul RE, Lambrechts L (2016) Determinants of Arbovirus Vertical Transmission in Mosquitoes. *PLOS Pathogens* 12, e1005548. https://doi.org/10.1371/journal.ppat.1005548

Li D, Liu C-M, Luo R, Sadakane K, Lam T-W (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. https://doi.org/10.14806/ej.17.1.200

Olmo RP, Todjro YMH, Aguiar ERGR, de Almeida JPP, Ferreira FV, Armache JN, de Faria IJS, Ferreira AGA, Amadou SCG, Silva ATS, de Souza KPR, Vilela APP, Babarit A, Tan CH, Diallo M, Gaye A, Paupy C, Obame-Nkoghe J, Visser TM, Koenraadt CJM, Wongsokarijo MA, Cruz ALC, Prieto MT, Parra MCP, Nogueira ML, Avelino-Silva V, Mota RN, Borges MAZ, Drumond BP, Kroon EG, Recker M, Sedda L, Marois E, Imler J-L, Marques JT (2023) Mosquito vector competence for dengue is modulated by insect-specific viruses. *Nature Microbiology* 8, 135–149. https://doi.org/10.1038/s41564-022-01289-4

Pan Y-F, Zhao H, Gou Q-Y, Shi P-B, Tian J-H, Feng Y, Li K, Yang W-H, Wu D, Tang G, Zhang B, Ren Z, Peng S, Luo G-Y, Le S-J, Xin G-Y, Wang J, Hou X, Peng M-W, Kong J-B, Chen X-X, Yang C-H, Mei S-Q, Liao Y-Q, Cheng J-X, Wang J, Chaolemen, Wu Y-H, Wang J-B, An T, Huang X, Eden J-S, Li J, Guo D, Liang G, Jin X, Holmes EC, Li B, Wang D, Li J, Wu W-C, Shi M (2024) Metagenomic analysis of individual mosquito viromes reveals the geographical patterns and drivers of viral diversity. *Nature Ecology & Evolution* 8, 947–959. https://doi.org/10.1038/s41559-024-02365-0

R Core Team (2023) *R: A Language and Environment for Statistical Computing*.

Sadeghi M, Altan E, Deng X, Barker CM, Fang Y, Coffey LL, Delwart E (2018) Virome of > 12 thousand *Culex* mosquitoes from throughout California. *Virology* 523, 74–88. https://doi.org/10.1016/j.virol.2018.07.029

Samy AM, Elaagip AH, Kenawy MA, Ayres CFJ, Peterson AT, Soliman DE (2016) Climate Change Influences on the Global Potential Distribution of the Mosquito *Culex quinquefasciatus*, Vector of West Nile Virus and Lymphatic Filariasis. *PLOS ONE* 11, e0163863. https://doi.org/10.1371/journal.pone.0163863

Sanborn MA, Klein TA, Kim H-C, Fung CK, Figueroa KL, Yang Y, Asafo-adjei EA, Jarman RG, Hang J (2019) Metagenomic Analysis Reveals Three Novel and Prevalent Mosquito Viruses from a Single Pool of *Aedes vexans nipponii* Collected in the Republic of Korea. *Viruses* 11, 222. https://doi.org/10.3390/v11030222

Siddell SG, Smith DB, Adriaenssens E, Alfenas-Zerbini P, Dutilh BE, Garcia ML, Junglen S, Krupovic M, Kuhn JH, Lambert AJ, Lefkowitz EJ, Łobocka M, Mushegian AR, Oksanen HM, Robertson DL, Rubino L, Sabanadzovic S, Simmonds P, Suzuki N, Van Doorslaer K, Vandamme A-M, Varsani A, Zerbini FM (2023) Virus taxonomy and the role of the International Committee on Taxonomy of Viruses (ICTV). *The Journal of General Virology* 104, 001840. https://doi.org/10.1099/jgv.0.001840

Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB, Davison AJ, Delwart E, Gorbalenya AE, Harrach B, Hull R, King AMQ, Koonin EV, Krupovic M, Kuhn JH, Lefkowitz EJ, Nibert ML, Orton R, Roossinck MJ, Sabanadzovic S, Sullivan MB, Suttle CA, Tesh RB, van der Vlugt RA, Varsani A, Zerbini FM (2017) Virus taxonomy in the age of metagenomics. *Nature Reviews Microbiology* 15, 161–168. https://doi.org/10.1038/nrmicro.2016.177

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stecher G, Tamura K, Kumar S (2020) Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Molecular Biology and Evolution* 37, 1237–1239. https://doi.org/10.1093/molbev/msz312

Thongsripong P, Chandler JA, Kittayapong P, Wilcox BA, Kapan DD, Bennett SN (2021) Metagenomic shotgun sequencing reveals host species as an important driver of virome composition in mosquitoes. *Scientific Reports* 11, 8448. https://doi.org/10.1038/s41598-021-87122-0

Valles SM, Chen Y, Firth AE, Guérin DMA, Hashimoto Y, Herrero S, de Miranda JR, Ryabov E, ICTV Report Consortium (2017) ICTV Virus Taxonomy Profile: *Iflaviridae*. *Journal of General Virology* 98, 527–528. https://doi.org/10.1099/jgv.0.000757

vegan: Community Ecology Package (2012).

Vincent M, Paty MC, Gerardin P, Balleydier E, Etienne A, Daoudi J, Thouillot F, Jaffar-Bandjee M-C, Team CI, Network L, Réunion on behalf of the study collaborators Réseau de médecins sentinelles de la, Menudier L (2023) From dengue outbreaks to endemicity: Reunion Island,

France, 2018 to 2021. *Eurosurveillance* 28, 2200769. https://doi.org/10.2807/1560-7917.ES.2023.28.29.2200769

Wei W, Shao D, Huang X, Li J, Chen H, Zhang Q, Zhang J (2006) The pathogenicity of mosquito densovirus (C6/36DNV) and its interaction with dengue virus type II in *Aedes albopictus*. *The American Journal of Tropical Medicine and Hygiene* 75, 1118–1126. https://doi.org/10.4269/ajtmh.2006.75.1118

Acknowledgements

We thank Yann Gomard, Cyrille Lebon and Sarah Hafsah for their help in collecting samples in the field.

Funding

This work was funded by the French ANSES PNR EST programme (project 2018/1/183, “DENSOTOOL," 2019–2023).

Conflict of interest disclosure

No financial conflicts of interest in relation to the content of the article.

Figures

Figure 1: Information on sampling locations on Reunion Island, Indian Ocean. Larvae from different breeding sites have been pooled by sampling location. In some cases breeding sites were hosting one or both of the species of interest.

Figure 2: **Viral communities recovery efforts. a**. Rarefaction curves of viral communities recovered from mosquito samples at the family level. **b.** Accumulation curve of viral communities recovered from mosquito samples at the family level.

Figure 3: Proportion of viral reads classified at the family level.
Aedes albopictus (red); *Culex quinquefasciatus* (blue).

Figure 4: Abundance of viruses found in *Aedes albopictus* (red) and *Culex quinquefasciatus* (blue). The horizontal axis represents the log (1+ number of reads attributed to each family). The number aside each dot represents the number of samples where the corresponding family was found.

Figure 5: Non-metric multi-dimensional scaling (MDS) plot of virus communities based on Bray-Curtis dissimilarity matrix, at the family level. A PERMANOVA, also based on Bray-Curtis dissimilarity matrix and permutational tests of dispersions (PERDISPs) was conducted in complement (R2 value: 17%), with a *p* value <0.0001. Samples are coloured by host species.

Figure 6: Heatmap representing the abundance of virus families in *Aedes albopictus* and *Culex quinquefasciatus* samples. The colours represent the normalised proportion of reads

attributed to each family. Sample names between '*' correspond to *Aedes albopictus* while others correspond to *Culex quinquefasciatus*.

Figure 7: Maximum likelihood phylogenetic tree based on the NS1 nucleotide sequences of 52 *Dipteran protoambidensovirus 1* (common virus name: Culex pipiens densovirus (CpDV)) sequences. The alignment of 892 nucleotides in length was produced by MAFFT v7.450 using the G-INS-i algorithm. The tree is mid-point rooted. Bootstrap values (100 replicates) are indicated at each node. Scale bar corresponds to nucleotide substitutions per site. The CpDV sequence obtained from our samples is coloured in red. CpDV clades were defined according to(Altinli et al., 2019).

Figure 8: Maximum likelihood phylogenetic tree based on the polyprotein amino acid sequences of 27 *Iflaviridae* sequences. The alignment of 2102 amino acids in length was produced by MUSCLE 5.1, and unconserved regions were trimmed by trimAL 1.4 using default parameters (Capella-Gutiérrez et *al.*, 2009). The tree is mid-point rooted. Bootstrap values (100 replicates) are indicated at each node. Scale bar corresponds to amino acid substitutions per site. The sequence obtained from our samples is coloured in red, and the blue sequences represented viruses isolated from dipteran species.

Figure 9: Maximum likelihood phylogenetic tree based on the major capsid protein amino acid sequences of 65 *Microviridae* sequences. The alignment of 487 amino acids in length was produced by MUSCLE 5.1, and unconserved regions were trimmed by trimAL 1.4 using default parameters (Capella-Gutiérrez et *al.*, 2009). The tree is mid-point rooted. Bootstrap values (100 replicates, ≥40) are indicated at each node. Scale bar corresponds to amino acid substitutions per site. The sequence obtained from our samples is coloured in red.

Table 1: Information on sample collection.

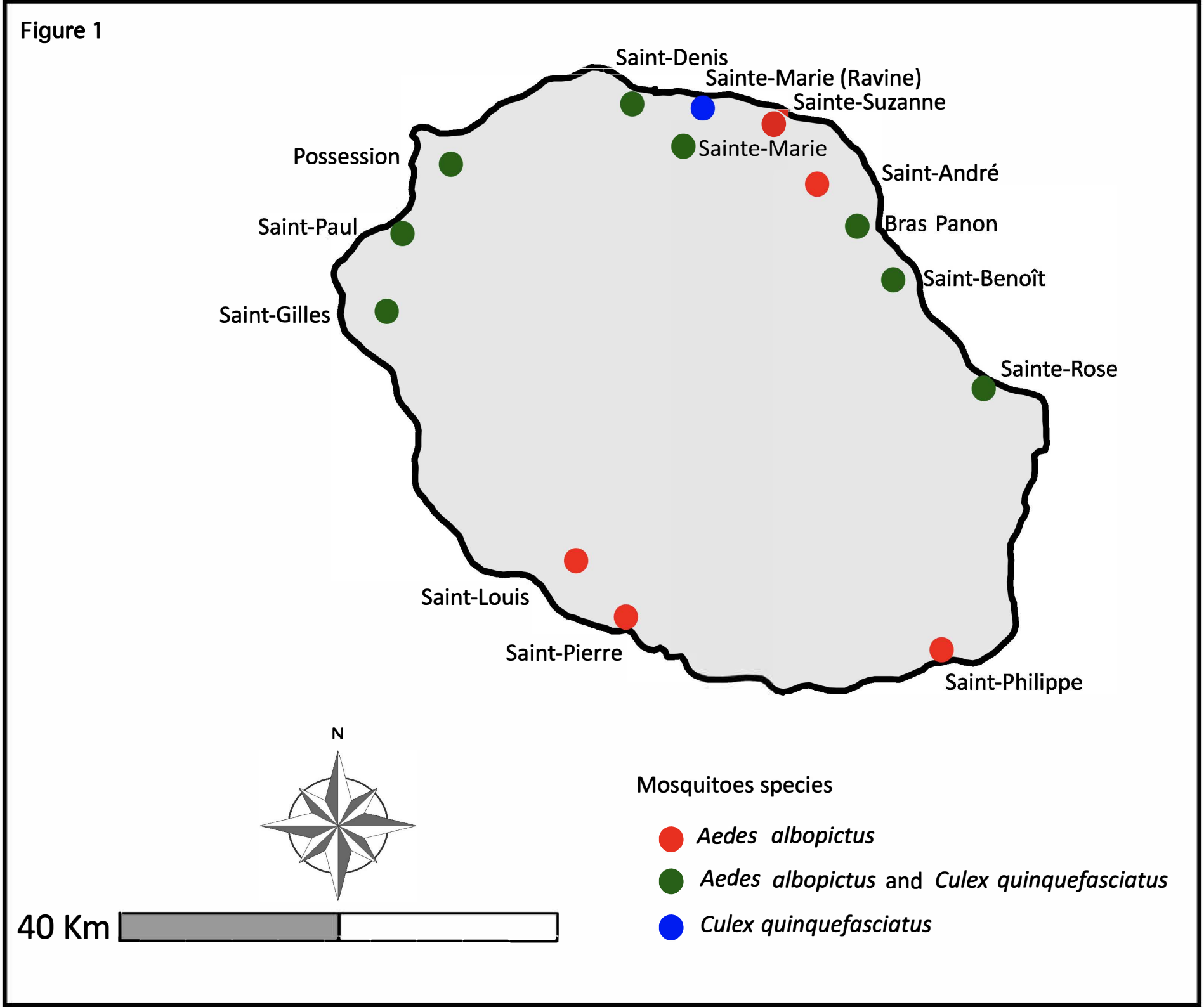
Table 2: Information on the five viral genomes reconstructed from mosquito viromic data.

Supplementary table 1: Viral genomes that were reconstructed from mosquito viromic data, including partial genomes that were not deposited into GenBank database.

Supplementary table 2: Contingency table which was used for statistical analyses.

Supplementary figure 1: Diversity analysis of *Aedes albopictus* and *Culex quinquefasciatus* viral communities at the family level. Boxplots indicate the observed numbers, Shannon diversity and Simpson diversity (left to right). Pairwise comparisons were conducted using Wilcoxon tests. The numbers indicate *p* values.

Figure 1


Saint-Denis
Sainte-Marie (Ravine)
Sainte-Suzanne
Sainte-Marie
Possession
Saint-André
Saint-Paul
Bras Panon
Saint-Benoît
Saint-Gilles
Sainte-Rose
Saint-Louis
Saint-Pierre
Saint-Philippe
N
40 Km
Mosquitoes species
Aedes albopictus
Aedes albopictus and Culex quinquefasciatus
Culex quinquefasciatus

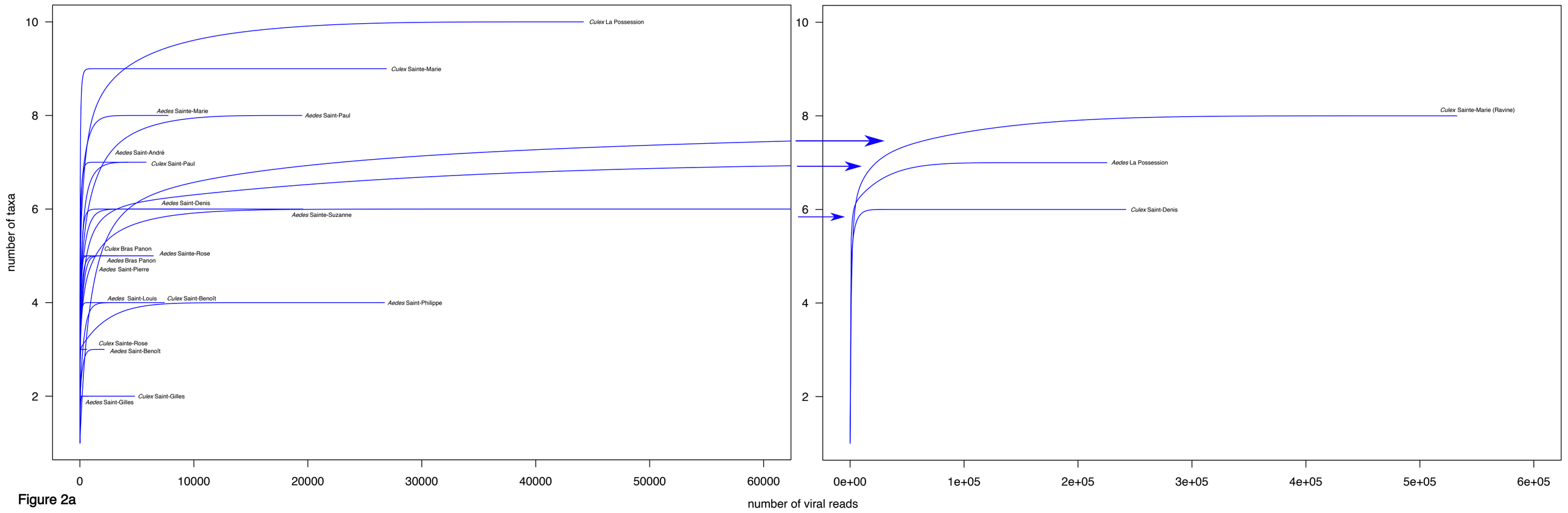

number of taxa
number of viral reads
Culex La Possession
Culex Sainte-Marie
Aedes Sainte-Marie
Aedes Saint-Paul
Aedes Saint-André
Culex Saint-Paul
Aedes Saint-Denis
Aedes Sainte-Suzanne
Culex Bras Panon
Aedes Sainte-Rose
Aedes Bras Panon
Aedes Saint-Pierre
Aedes Saint-Louis
Culex Saint-Benoît
Aedes Saint-Philippe
Culex Sainte-Rose
Aedes Saint-Benoît
Culex Saint-Gilles
Aedes Saint-Gilles
Culex Sainte-Marie (Ravine)
Aedes La Possession
Culex Saint-Denis

Figure 2a

Figure 2b

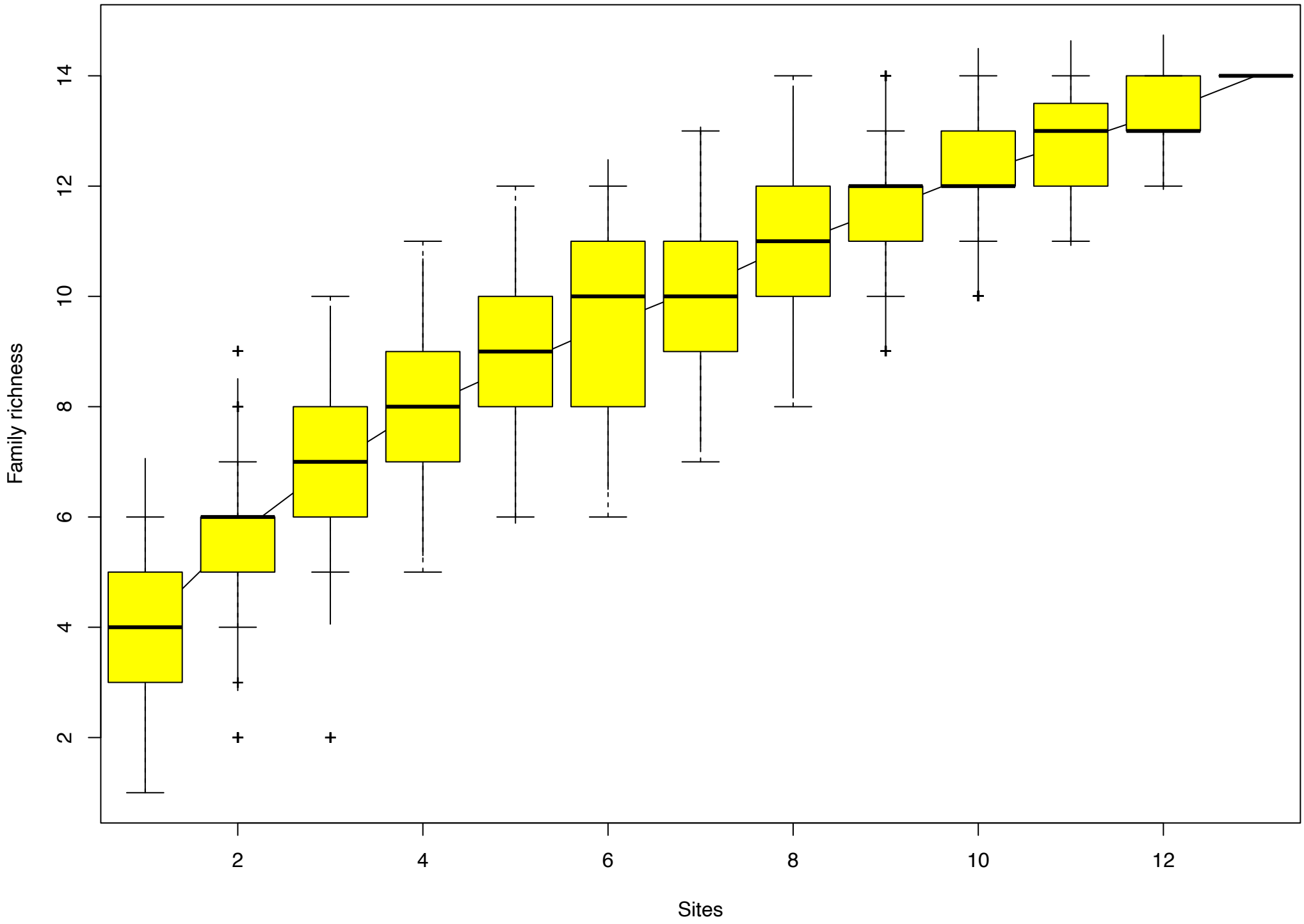


Figure 3

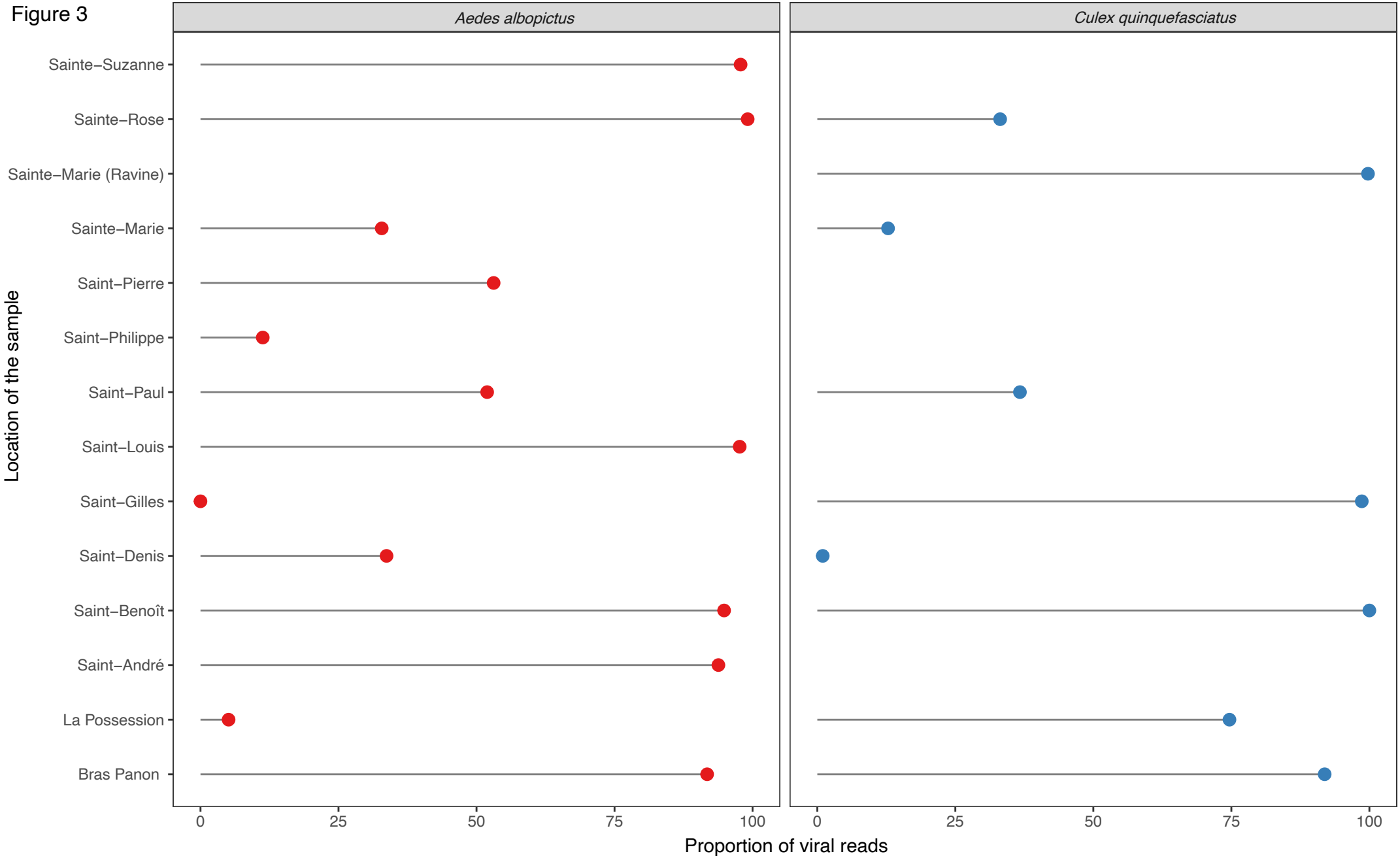

Aedes albopictus
Culex quinquefasciatus
Sainte−Suzanne
Sainte−Rose
Sainte−Marie (Ravine)
Sainte−Marie
Saint−Pierre
Saint−Philippe
Saint−Paul
Saint−Louis
Saint−Gilles
Saint−Denis
Saint−Benoît
Saint−André
La Possession
Bras Panon
Location of the sample
0
25
50
75
100
Proportion of viral reads

Figure 4

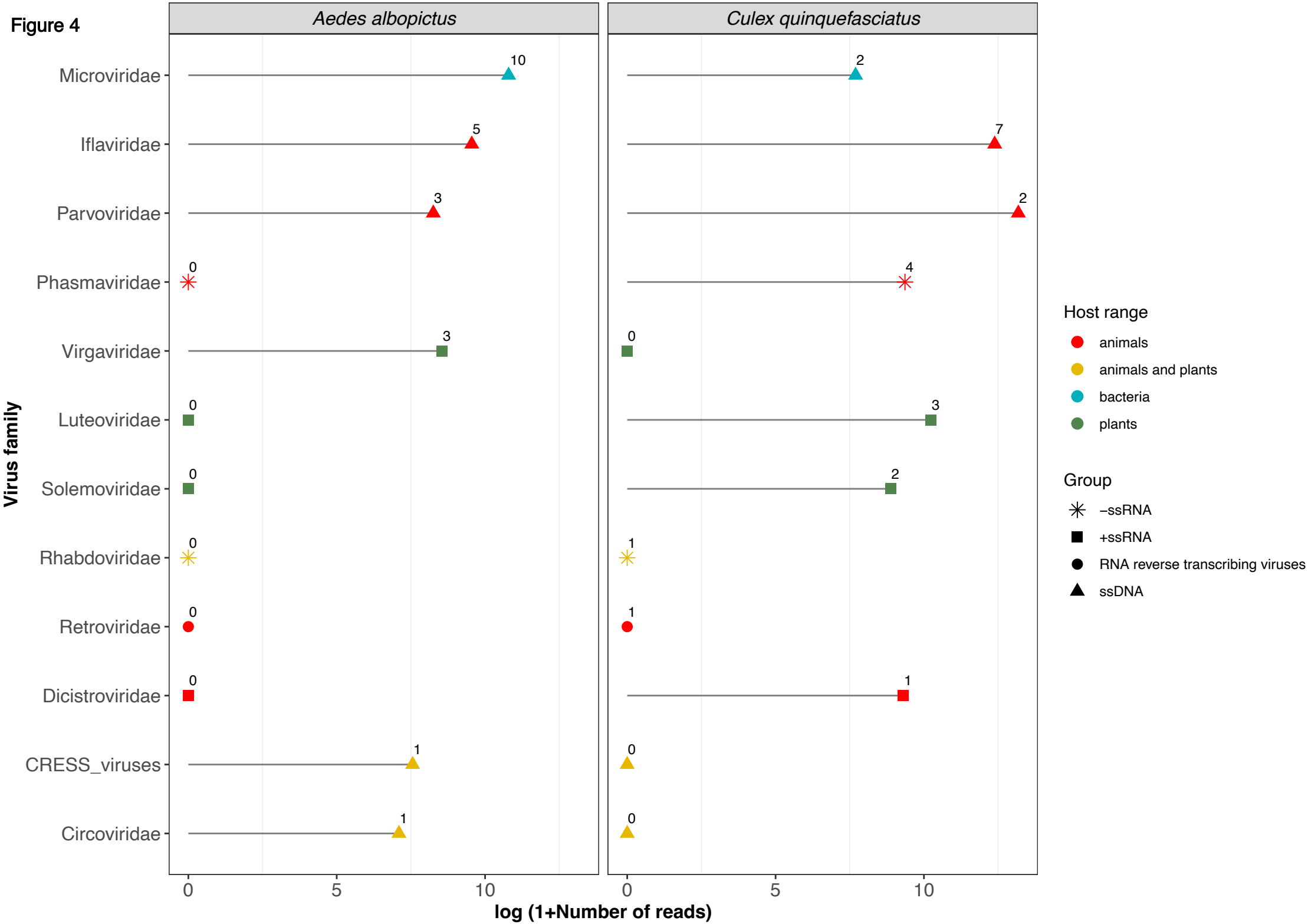

Aedes albopictus
Culex quinquefasciatus
Virus family
Microviridae
Iflaviridae
Parvoviridae
Phasmaviridae
Virgaviridae
Luteoviridae
Solemoviridae
Rhabdoviridae
Retroviridae
Dicistroviridae
CRESS_viruses
Circoviridae
10
5
3
0
3
0
0
0
0
0
1
1
2
7
2
4
0
3
2
1
1
1
0
0
0
5
10
log (1+Number of reads)
Host range
animals
animals and plants
bacteria
plants
Group
−ssRNA
+ssRNA
RNA reverse transcribing viruses
ssDNA

Figure 5

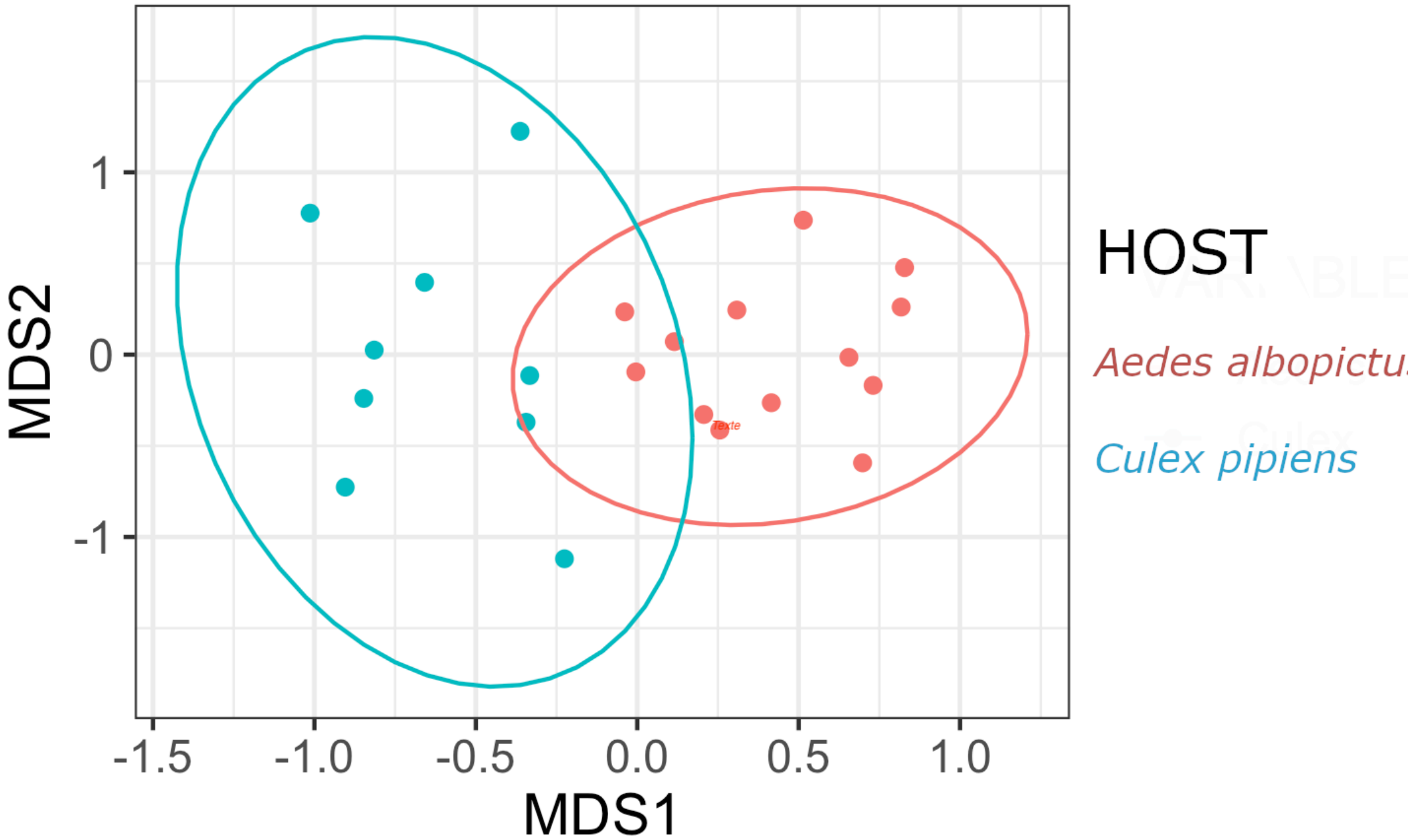


Figure 6

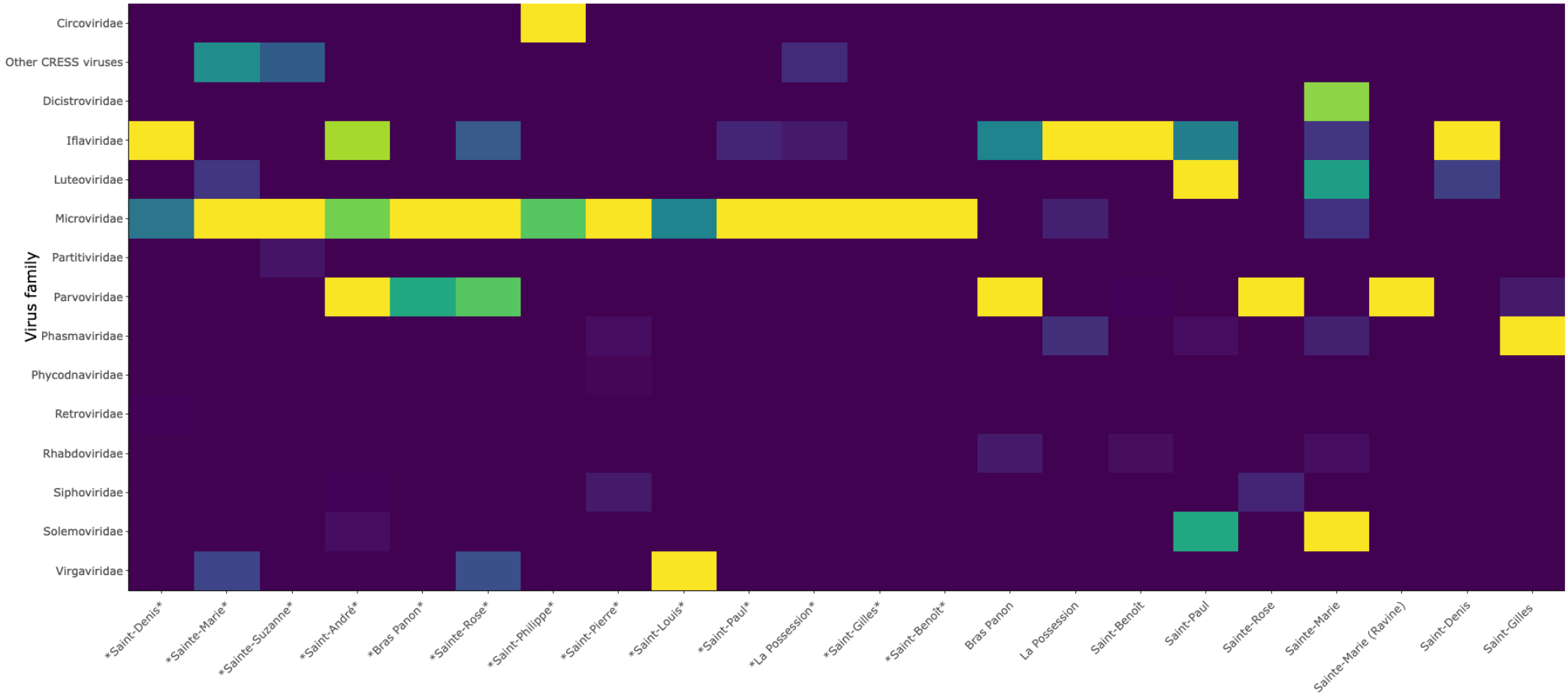
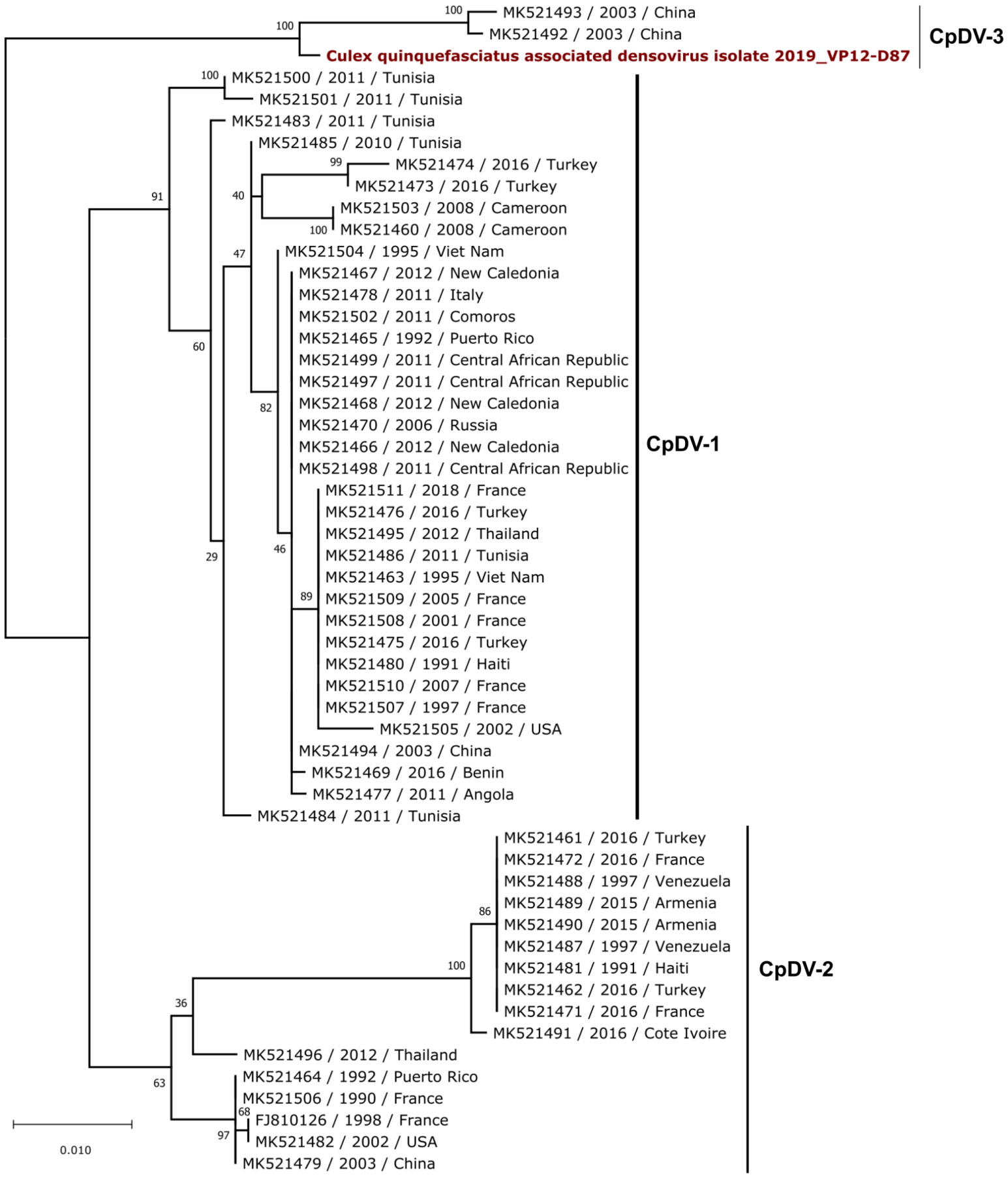


Figure 7



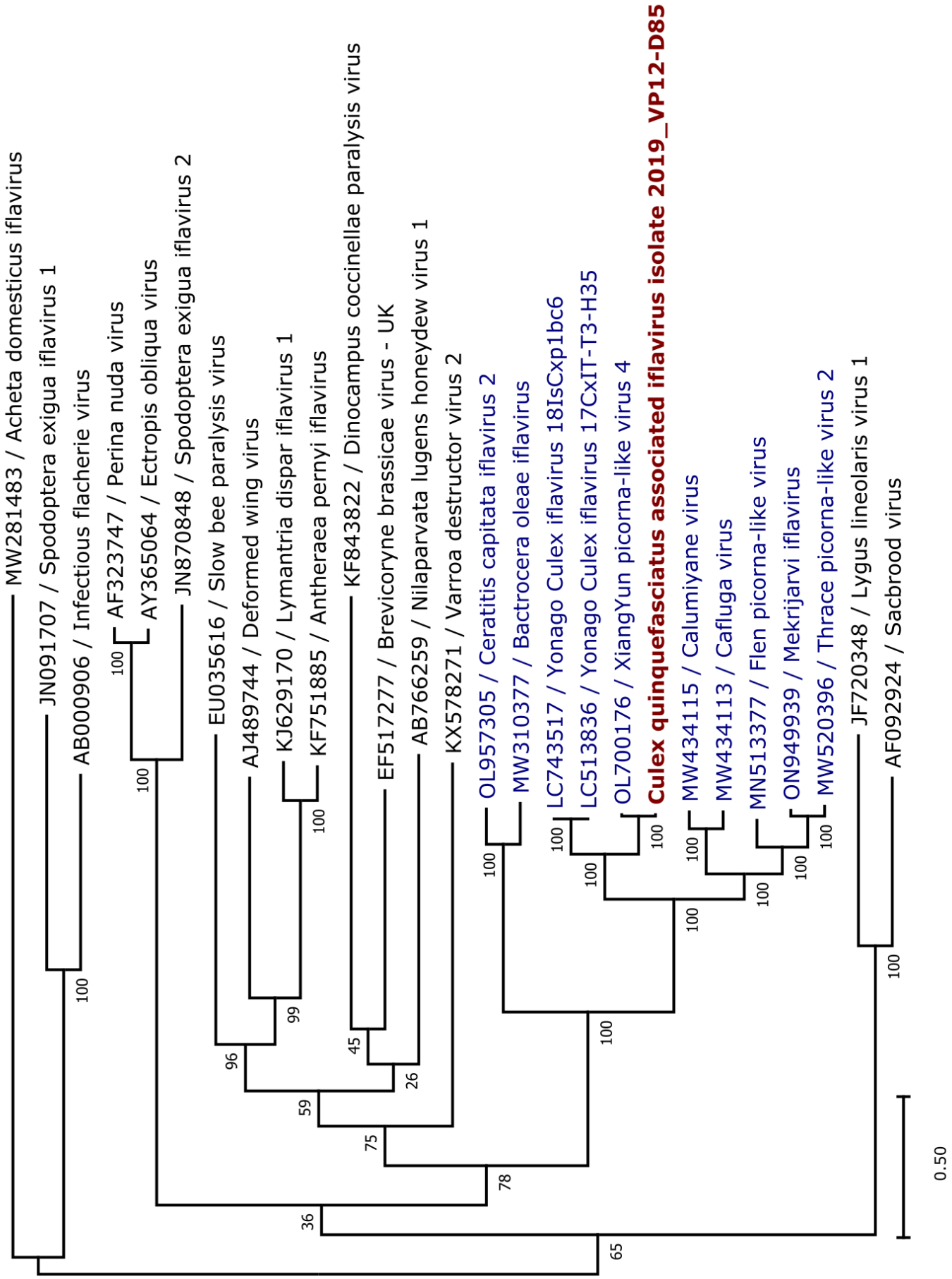


Figure 8

Figure 9