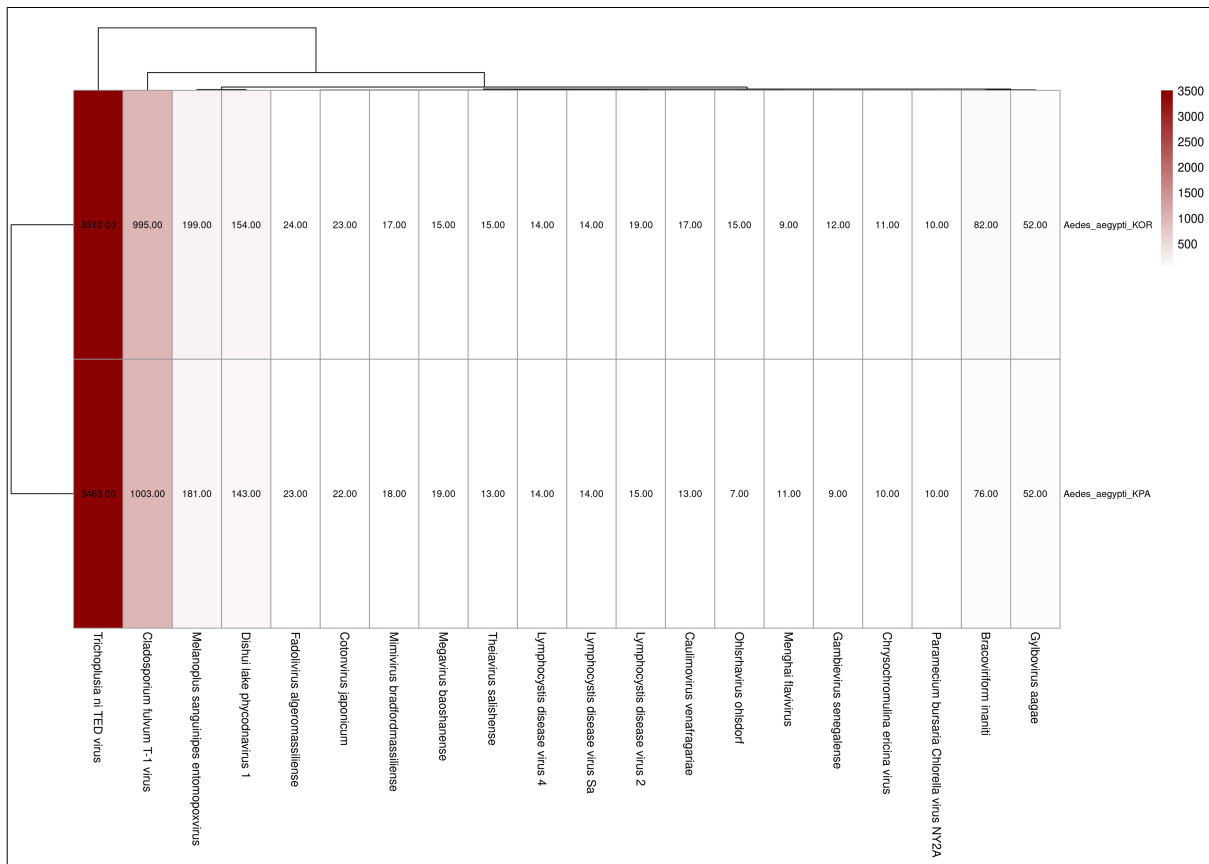
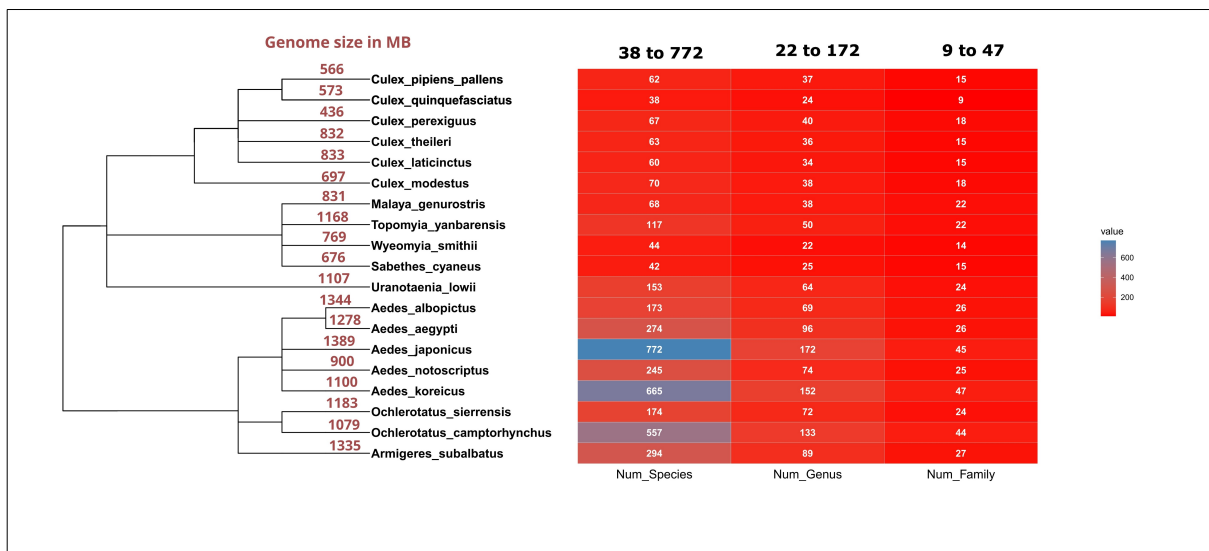


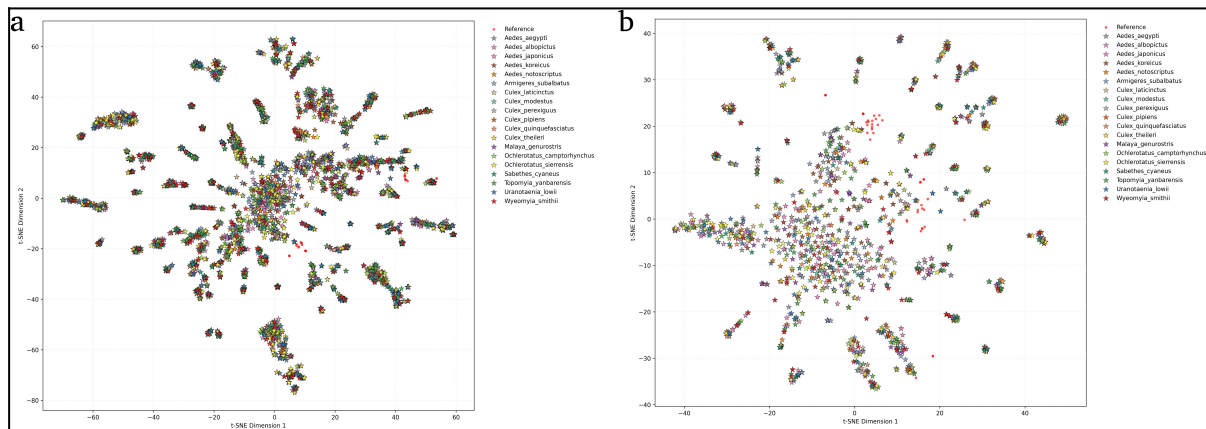


S. Figure 1: Abundance of the top 20 EVEs in two Indian *Aedes aegypti* genomes. The most prevalent viral elements shows global mosquito profiles.

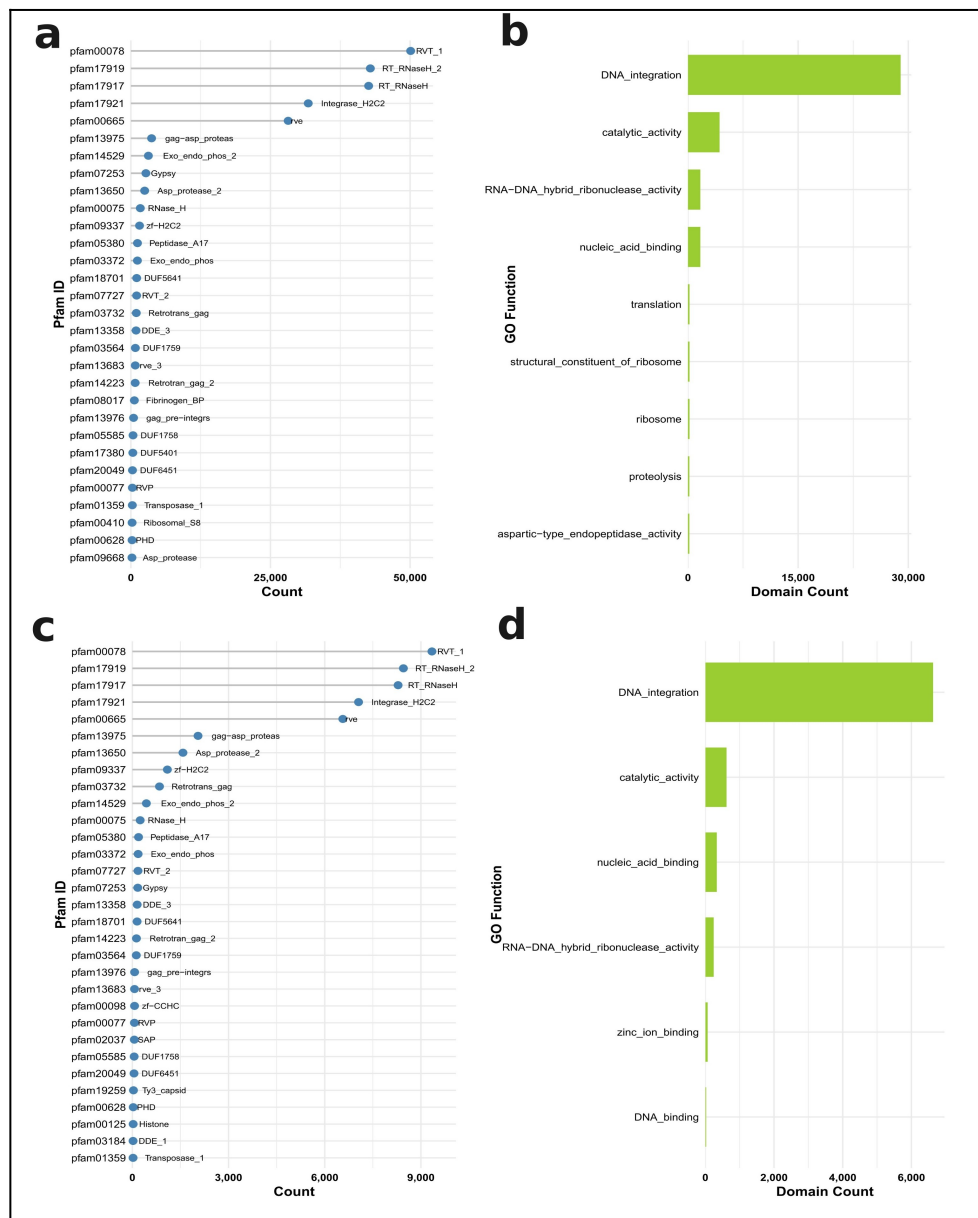


S. Figure 2: Taxonomic diversity of endogenous viral elements (EVEs) across *Culicinae* mosquito genomes. The heatmap displays the number of unique viral species, genera, and families represented by EVEs in nineteen mosquito genomes, ordered according to NCBI taxonomy-based relationships generated with phyloT (<https://phylot.biobyte.de/>). Numbers within each cell denote the total count of viral taxa detected at the corresponding taxonomic

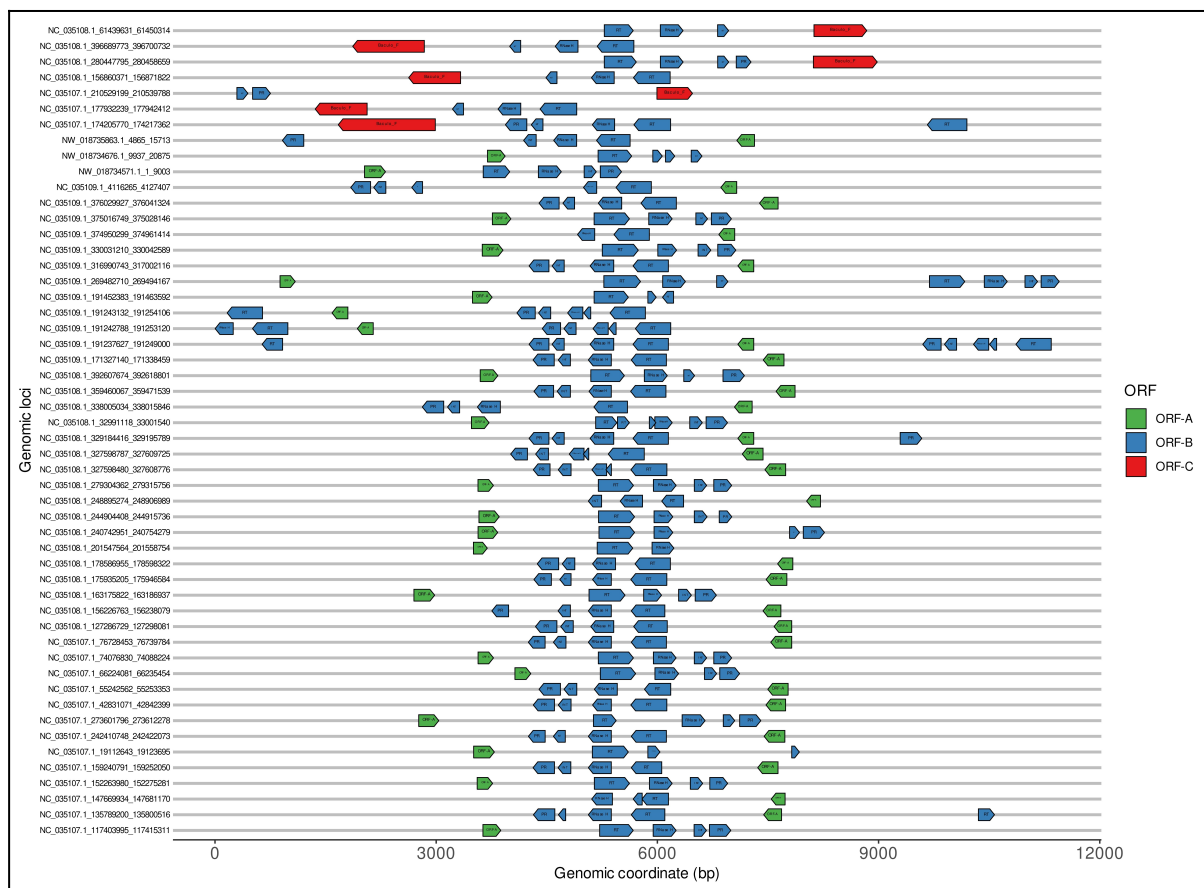
rank, with color intensity indicating relative diversity. Genome assembly sizes (Mb) for each species are provided alongside species names.



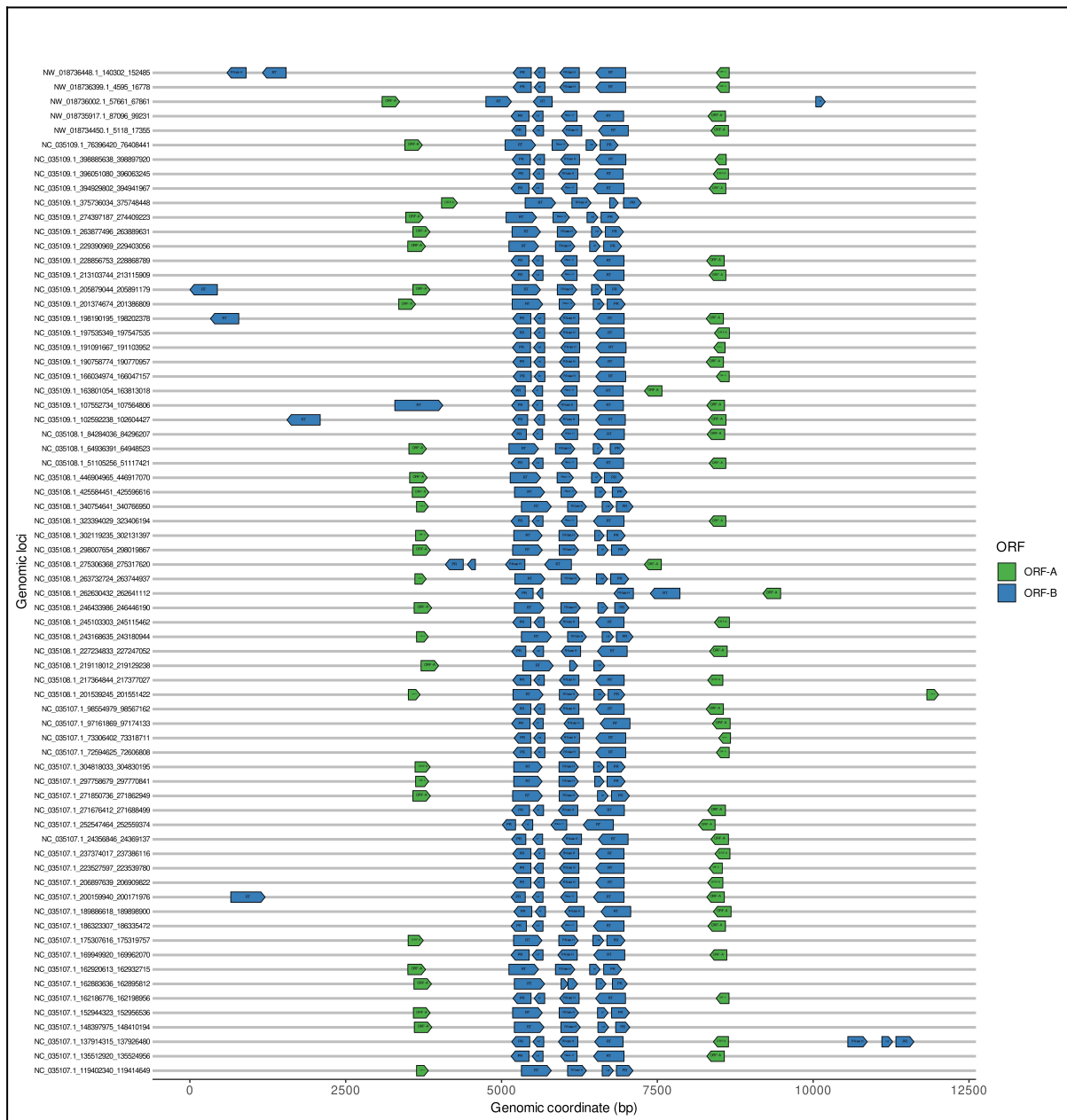
S. Figure 3: Clustering of representative TED-like and Clad-like endogenous viral element (EVE) proteins and reference viral proteins using DBSCAN/t-SNE. Protein sequences were projected into two-dimensional space via t-SNE and clustered with DBSCAN. Colored stars indicate EVEs from various mosquito species, while red circles denote reference viral proteins. **(A)** In the *TED*-like dataset, 3,516 EVEs and 24 reference proteins formed 53 clusters. **(B)** In the *Clado*-like dataset, DBSCAN resolved 1,280 protein sequences (1,236 EVEs and 44 reference proteins) into 40 groups (39 multi-sequence clusters and one singleton group).



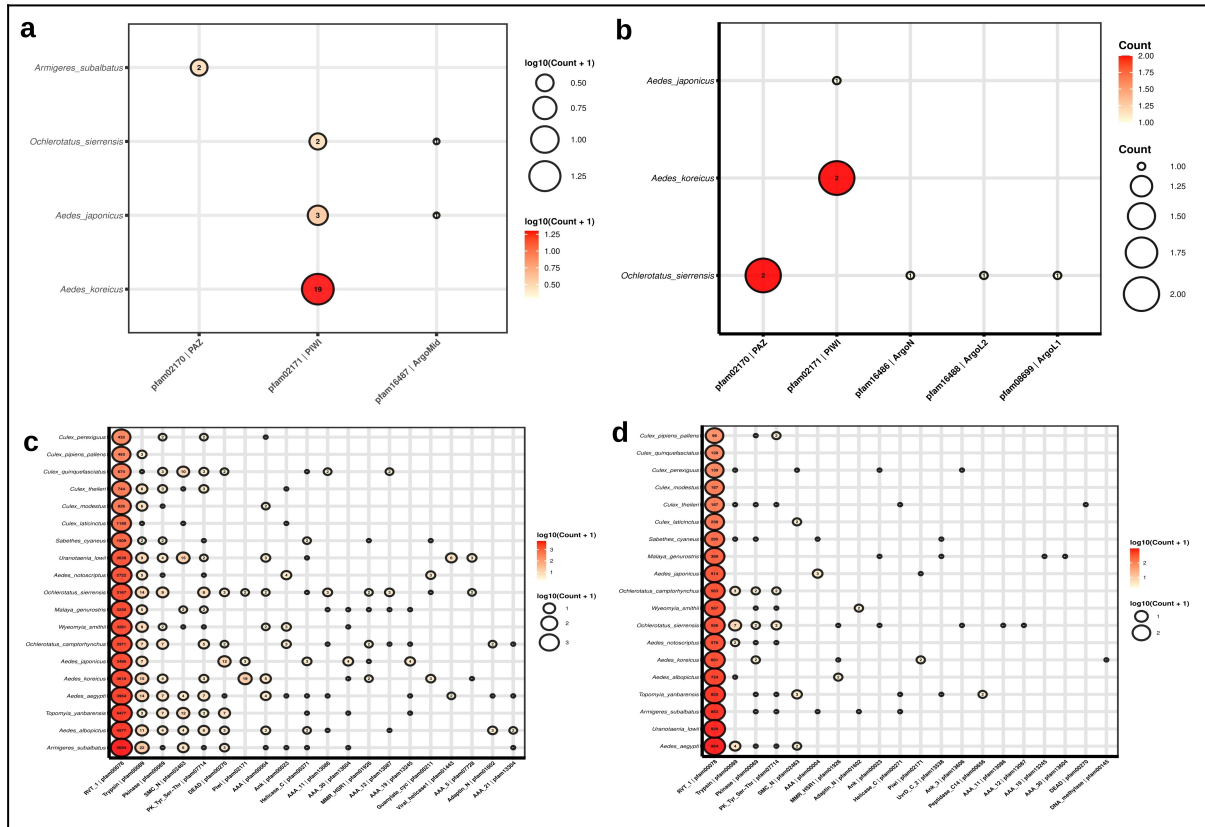
S. Figure 4: Pfam protein domain in the 5kb flanking region of *TED*-like ORF B & *Clado*-like RT. (A) The top 30 most frequently detected Pfam domains are represented. (B) Pfam-to-GO functional enrichment in the flanking region. (C) Pfam-rich protein domain in 5Kb flank (D) Pfam-to-GO functional enrichment in 5Kb flank region.



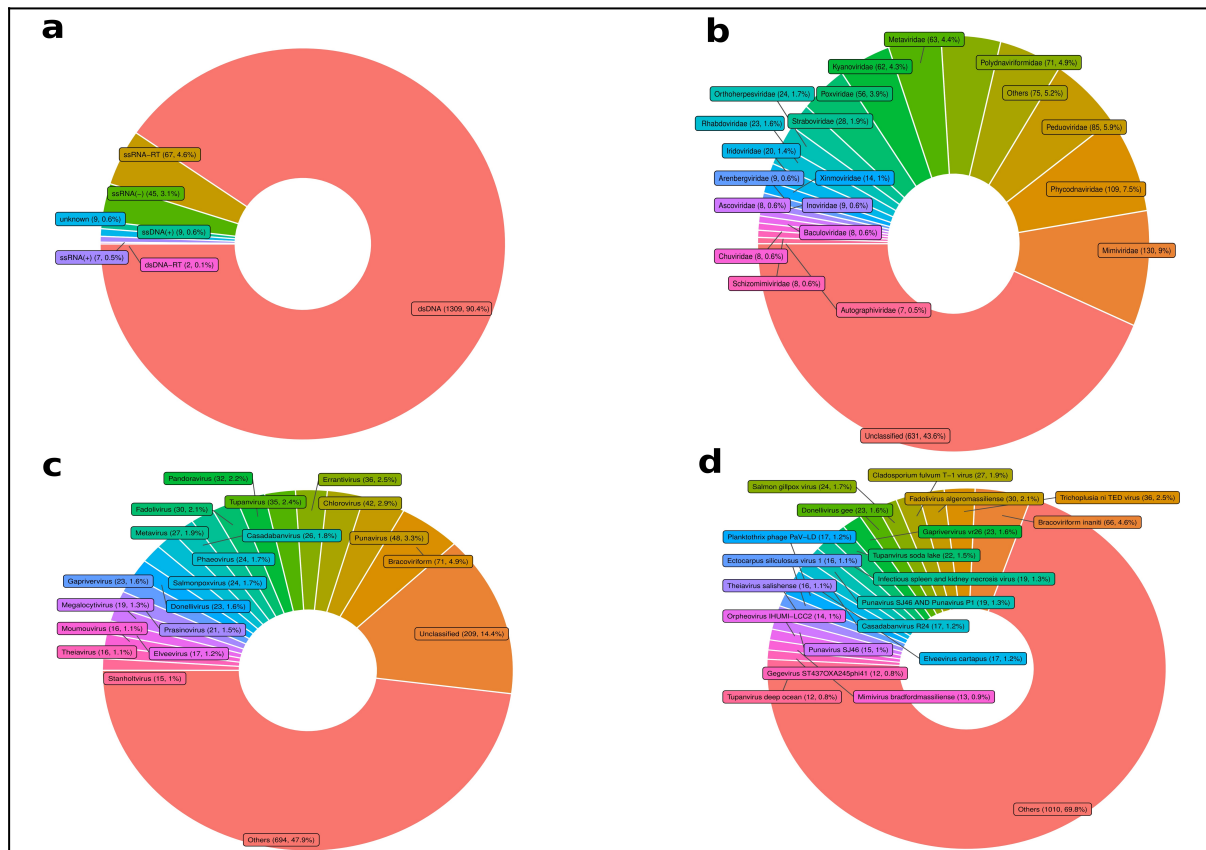
S. Figure 5: Pfam domain synteny of TED-like elements identified in *Aedes aegypti*. Each horizontal track represents an individual genomic locus containing a TED-like element. ORF-A (green) encodes the **pfam03732** (retrotransposon Gag) domain, ORF-B (blue) contains the conserved Pol-associated domains **pfam00078** (Reverse Transcriptase), **pfam00665** (Integrase core domain), **pfam17917** (RNase H), and **pfam17921** (Integrase zinc-binding domain), while ORF-C (red) encodes the **pfam12259** (Baculo_F protein) domain. Arrows indicate ORF orientation, and domain positions are shown relative to their genomic coordinates within each locus.



S. Figure 6: Pfam domain synteny of the *Clado*-like element in *Aedes aegypti*. Each horizontal track represents an individual genomic locus containing a Clado-like element. ORF-A (green) encodes a pfam03732 (retrotransposon Gag) domain, whereas ORF-B (blue) harbors a conserved complement of Pol-associated domains, including pfam00078 (Reverse Transcriptase), pfam00665 (Integrase core domain), pfam17917 (RNase H), and pfam17921 (Integrase zinc-binding domain). Arrows indicate the predicted orientation of each ORF, and domain positions are shown relative to their genomic coordinates within each locus.



S. Figure 7. Distribution of PIWI-associated and prokaryotic defense-related protein domains within ± 5 kb flanking regions of *TED*-like and *Clado*-like EVEs in mosquito genomes. (A) Bubble plot illustrating the distribution and abundance of canonical PIWI-associated domains detected within ± 5 kb flanking regions of *TED*-like EVEs. Bubble size and color intensity represent the frequency of PIWI-related protein domains across mosquito genomes, with numbers inside bubbles indicating raw domain counts. **(B)** Bubble plot showing the distribution of PIWI-associated domains identified within ± 5 kb flanking regions of *Clado*-like EVEs. Bubble size and color intensity indicate the frequency of PIWI-related protein domains across mosquito species. **(C)** Bubble plot depicting prokaryotic defense-related protein domains identified within ± 5 kb flanking regions of *TED*-like EVEs. The top 20 most abundant defense-associated domains detected across mosquito genomes are shown. Bubble size and color intensity correspond to domain abundance, with numbers inside bubbles representing raw counts. **(D)** Bubble plot showing the distribution of prokaryotic defense-related domains detected within ± 5 kb flanking regions of *Clado*-like EVEs. The top 20 most abundant defense-associated domains identified across mosquito genomes are presented.



S. Figure 8: Diversity of putative EVEs identified in assembled mosquito transcriptomes. (A) Molecular type of putative EVEs (B), Diversity at family level (C), Diversity at Genus level (D), Diversity at Species level.