Viral metagenome characterization reveals species-specific virome profiles in Triatominae populations from the southern United States

Insan Habib[1¶], Juan C. Hernandez-Valencia [1¶], Jana Martinu[1], Eva Nováková[1,2*]

[1]University of South Bohemia, Faculty of Science, Ceske Budejovice, Czech Republic

[2]Biology Centre of the Czech Academy of Sciences, Institute of Parasitology, Ceske Budejovice, Czech Republic

[*]Corresponding author

E-mail: novake01@prf.jcu.cz

¶ JCH and IH contributed equally to this work.

Key words

Triatominae; Kissing bugs; *Trypanosoma cruzi*; Chagas disease; metatranscriptomics; virome; Insect Viruses.

Abstract

Kissing bugs (Triatominae) are hematophagous insects and the principal vectors of *Trypanosoma cruzi*, the causative agent of Chagas disease. While their bacterial microbiomes have received considerable attention, the diversity of viruses associated with these insects remains poorly understood. To address this gap, we investigated the metavirome of five Triatominae species from the southern United States (*Triatoma rubida, T. sanguisuga, T. gerstaeckeri, T. indictiva,* and *Hospesneotomae protracta*), sampled in Texas, New Mexico, and Arizona. We sequenced gut and reproductive tissues from 23 field-collected specimens and assembled 41 viral operational taxonomic units (vOTUs), 40 of which are novel and together constitute 13 viral families, including *Chuviridae, Arenaviridae, Orthomyxoviridae, Partitiviridae, Solemoviridae, Circoviridae, Rhabdoviridae, Microviridae, Xinmoviridae, Astroviridae, Narnaviridae, Tombusviridae,* and the order Elliovirales. The vOTUs composition and abundance analysis examined variables including host species, sex, tissue type, blood meal, and *T. cruzi* infection status, showing that metavirome diversity varied significantly among host species. Our findings demonstrate a species-specific metavirome and the presence of virus taxa linked to insects, plants, and vertebrates, highlighting the complex ecological interactions between viruses and triatomines. This study uncovers a diverse and largely novel set of metaviromes within North American Triatominae, providing a foundation for future research on virus-vector interactions.

36 # Introduction

37 Metagenomic and metatranscriptomic surveys in recent years have vastly expanded our
38 understanding of the viromes associated with blood-feeding arthropods. These studies reveal that
39 hematophagous vectors harbor a surprising diversity of viruses, ranging from well-known
40 arboviruses to numerous insect-specific viruses (ISVs) that do not infect vertebrates (1). While
41 mosquitoes and ticks have received the most attention of vector virome research (2-5), kissing
42 bugs (*Triatominae*) and their associated viruses have remained largely understudied. Despite the
43 well-established role of *Triatominae* as vectors of *Trypanosoma cruzi*, the etiological agent of
44 Chagas disease, current knowledge of their virome diversity, distribution, ecological dynamics,
45 and host interactions remains limited (6). Until recently, only eight viruses had been described
46 from *Triatominae*: the dicistrovirus Triatoma virus (TrV) and seven Rhodnius prolixus virus
47 strains (RpV1-7). TrV, a non-enveloped positive-sense single-stranded RNA virus from the
48 *Dicistroviridae* family, was first identified in *Triatoma infestans* populations in Argentina (7, 8).
49 Subsequent studies demonstrated that TrV is the only known triatomine viral pathogen (9-12). Due
50 to its association with high nymphal mortality, delayed development, and impaired oogenesis, TrV
51 has been proposed as a potential biological control agent for Chagas disease vectors (8, 13). In
52 contrast, the phenotypic effects of the RpV1-7 viruses remain unknown. Discovered through
53 transcriptome sequencing of *Rhodnius prolixus* tissues (14), these viruses are evolutionarily
54 distinct from TrV and belong to the families *Iflaviridae* (RpV1, RpV2), *Permutotetraviridae*
55 (RpV3, RpV4, RpV7), and *Solemoviridae* (RpV5, RpV6). Notably, both TrV and RpVs appear to
56 be vertically transmitted to progeny (8, 14), suggesting persistent infections within triatomine
57 populations.

58 The first virome oriented surveys in *Triatominae* were conducted by Brito et al. (14) and Bourke
59 et al. (15), who used RNA-seq data from public repositories to uncover a broader diversity of
60 viruses associated with kissing bugs. Mining the datasets from 122 *Triatominae* specimens (eight
61 species from six countries), the later study identified six distinct viruses, including the previously

62 known RpV4-6 and widespread detection of TrV in nearly half of all screened individuals (15).
63 This work notably expanded TrV's known host range, reporting its presence for the first time
64 in *Triatoma brasiliensis* and *Mepraia* species. Additionally, two novel viruses were detected in
65 kissing bugs: Meccus longipennis virus 1 and Drosophila melanogaster Nora virus. The former
66 was identified in multiple *T. infestans* populations across Argentina, Brazil, Chile, and Peru, while
67 the latter (a picorna-like virus previously known from fruit flies) was found in *T. infestans* in
68 Argentina. These findings emphasize the vast, yet underutilized, potential of publicly available
69 datasets for identifying both known and novel viruses, while also revealing significant gaps in our
70 understanding of Triatominae-associated viromes.

71 The present metatranscriptomic study focuses on five *Triatominae* species native to the southern
72 United States and provides a comprehensive analysis of the associated viral metagenomes. By
73 characterizing viral communities in both gut and reproductive tissues of *Triatoma*
74 *rubida*, *Triatoma gerstaeckeri*, *Hospesneotomae protracta*, *Triatoma sanguisuga*, and *Triatoma*
75 *indictiva*, we aim to gain insight into tissue-specific viral diversity and transmission potential. In
76 addition, we investigate how ecological and biological factors, such as *T. cruzi* infection status and
77 the identity of recent blood meals, influence virome composition and structure across hosts and
78 species. Identifying novel viruses associated with these medically relevant insects may reveal
79 candidates for biological control, as seen with TrV adversely affecting triatomine development
80 and survival.

81

Materials and methods

Sample set

The sample set was designed to include digestive and reproductive tracts from fifth instars, males, and females of five Triatominae species sampled from Arizona, New Mexico, and Texas in 2022. The individuals retrieved by active searches through host nests were morphologically determined to species, preserved in RNA later and stored in -20°C before dissection. Whole abdomens were used for 2 females, reproductive tissues and digestive tracts (excluding crop) were dissected from 11 individuals (5 females, 4 males, and 2 L5s) using a Nikon SMZ745T stereomicroscope and preserved in liquid nitrogen (S1 Table).

Total DNA/RNA extraction, Next-generation sequencing, and blood meal analysis

DNA and RNA templates from individual tissues were isolated with AllPrep DNA/RNA Micro Kit (Qiagen) following the manufacturer´s protocol. The quality of RNA was checked with Agilent 2100 Bioanalyzer (Agilent Technologies), and the sample concentration was measured using Qubit RNA High Sensitivity Assay Kit (Thermo Fisher Scientific). The samples passing QC (n=23, S1 Table) were used for cDNA libraries construction, and sequencing of the 150bp long PE reads on Illumina NovaSeq 6000 was carried out by Novogene (United Kingdom). Blood meal analysis was conducted with 12 DNA isolates derived from gut tissues (S2 Table). For the identification of bloodmeal source, we used vertebrate universal primers, amplifying 215 bp long fragment of the 12S rRNA gene: L1085 (5'-CCCAAACTGGGATTAGATACCC-3') and H1259 (5'-GTTTGCTGAAGATGGCGGTA-3') (Kitano et al. 2007). PCR products were purified and sent

104 for Sanger sequencing. Identity of the sequences was assessed based on the first BLASTN hits
105 against the core_nt database.

106 Raw data processing, assembly, and viral metagenome identification

107 Quality control of 23 raw sequence datasets was performed using the Trimmomatic (v0.39) tool
108 with default parameters to remove the low-quality reads and overrepresented sequences (16).
109 Clean reads were used as input and *de novo* assembled into contigs using MetaSPAdes (v3.15.2)
110 (K-mers= 21, 33, 55; other parameters were kept at default) (17). All the assembled contigs were
111 aligned to the protein version (v26.0) of the Reference Viral Database (RVDB) (18) using the
112 DIAMOND (v v0.9.14.115) with BLASTX algorithm using the --very-sensitive option, E-value
113 of 1e-3 and the default parameters (19). Additionally, to detect additional potential RNA-
114 dependent RNA Polymerase (RdRp) from RNA viruses, publicly available HMM profiles of
115 different RdRp of eukaryotic RNA virus families were obtained from Pfam database v36.0
116 (Mononega- and Chuviridae-like: Mononeg_RNA_pol [PF00946], Reoviridae-like: RdRP_5
117 [PF07925], Flaviviridae-like: Flavi_NS5 [PF00972], Bunyavirales-like: Bunya_RdRp [PF04196],
118 Narnaviridae-like: Mitovir_RNA_pol [PF05919], Nidovirales-like: RdRP_1 [PF00680],
119 Tymovirales-like and Hepe-Virga-like: RdRP_2 [PF00978], Tombusviridae-like and Nodaviridae-
120 like: RdRP_3 [PF00998], Toti-, Luteo-, and Sobemoviridae-like: RdRP_4 [PF02123],
121 Viral_RdRp_C [PF17501], and Birnaviridae-like: Birna_RdRp [PF04197]), available through
122 InterPro (20, 21). From all the retrieved profile HMMs a combined HMM profile database was
123 built using *hmmbuild* from the HMMER suite (22). The contigs were translated to six possible
124 open reading frames (ORFs) obtained from the online VirtualRibosome tool
125 (https://services.healthtech.dtu.dk/services/VirtualRibosome-2.0/) (23), and the amino acid (aa)
126 sequences were used as an input for *hmm*search. To avoid false positives, the contigs with virus
127 hits in RVDB and HMM were aligned against the NCBI's non-redundant (nr) protein database via
128 NCBI-BLAST with an E-value of 1e-5 and the default parameters (24). Furthermore, the non-
129 redundant contigs greater than 300bp (nt) were selected for functional and structural annotation,

130 and phylogenetic analyses. The virus sequences were taxonomically assigned following the
131 guidelines for sequence-based taxonomy provided by the International Committee on Taxonomy
132 of Viruses (ICTV) (25). The newly discovered virus sequences were named based on the genus or
133 subfamily and the species of the host, or the abbreviation "spp." if found in multiple species,
134 followed by the standard virus taxon according to their taxonomic classification.

135 Molecular confirmation of Triatominae species and host 136 phylogenetic background for the virome analyses

137 Molecular confirmation of the involved species, especially *T. sanguisuga* and *T. indictiva* was
138 performed by recovering 8 mitochondrial genes, including, Cytochrome C Oxidase I (COXI),
139 Cytochrome C Oxidase II (COXII), Cytochrome C Oxidase III (COXIII), NADH dehydrogenase
140 subunit 1 (ND1), NADH dehydrogenase subunit 4 (ND4), NADH dehydrogenase subunit 5 (ND5),
141 cytochrome b (cytb) and 16S ribosomal RNA (16S rRNA) sequences, from each meta-assembly.
142 The gene sequences were aligned separately using MAFFT algorithm implemented in Geneious
143 Prime v.2022.0.2 (http://www.geneious.com/) and then concatenated into the final matrix
144 containing 9032 nucleotides. The phylogenetic analysis was conducted with a maximum
145 likelihood approach using PhyML. Bootstrap analysis was run with 100 iterations.

146 Trypanosoma cruzi meta-assembly screening

147 To investigate viral metagenome diversity in the context of *Trypanosoma cruzi* infection, we
148 explored our data to identify infected and uninfected *Triatoma* and *Hospesneotomae* individuals.
149 All the assembled contigs were first subjected to BLASTN searches against the custom database
150 created from all available *T. cruzi* genomes from TriTryDB (https://tritrypdb.org/) (26). Identity
151 of all contigs producing an alignment of 200bp and longer in the first search was confirmed by
152 BLASTN against NCBI's nt database (S1 Table).

153 ## Phylogenetic analysis and annotation of viral sequences

154 For structural and functional annotation of viral sequences, the amino acid sequence of identified
155 ORFs was queried against the Protein Data Bank (PDB) database using HHpred method (27). For
156 the phylogenetic analysis of the assembled virus sequences, the amino acid sequences of the RdRp
157 were mainly used since they are highly conserved in nature. Where available, we also used the
158 nucleocapsid, and capsid sequences. The datasets were assembled from all BLASTP hits with more
159 than 30% identity found in the nr database. The host and geographic metadata for all the hits were
160 retrieved using the NCBI Entrez-direct tool (v19.4) (28) with a custom bash script and added to
161 the header of each FASTA file. To reconstruct phylogenetic relationships among identified virus
162 sequences, all the datasets were aligned using MAFFT under automatic settings (29). Aligned
163 sequences were trimmed using trimAl v1.4 using the *-gappyout* algorithm to remove poorly
164 aligned positions (30). Maximum-likelihood trees were inferred in IQ-TREEE2, where
165 *ModelFinder Plus* with partition merging (MFP+MERGE) selected the optimal amino-acid
166 substitution model (31). Node support was evaluated with 1000 ultrafast bootstrap replicates and
167 1000 SH-aLRT tests.
168

169 ## Viral metagenome diversity and abundance in Triatominae

170 Diversity analyses were conducted with an abundance-based matrix built using read counts
171 assigned to identified viral non-redundant contigs, hereafter referred to as viral Operational
172 Taxonomic Units (vOTUs). The read counts were obtained by mapping quality-filtered reads to
173 the vOTUs collection through the BBmap tool (v39.01). To avoid false positives, we used a
174 minimum threshold of 99% nucleotide identity for read mapping and >75% read coverage across
175 the entire length of each vOTU (32). Subsequently, the abundance matrix was normalized based
176 on sequence length and library size (quality-filtered reads per sample). This normalization was
177 estimated as Reads Per Kilobase per Million mapped reads (RPKM).

178 The diversity analyses considered four factors: 1) species (*T. rubida*, *T. sanguisuga, T. indictiva*,
179 *T. gersteackeri*, and *H. protracta*), 2) tissue type (ovaries, testes, gut, and complete abdomen), 3)
180 geographical location (Arizona, New Mexico, and Texas), and 4) sex (male and female). Alpha
181 diversity was evaluated using the Richness (S), Shannon-Weaver (H), Simpson (1-D), and
182 Evenness (E) indices, while Beta diversity was calculated using Bray-Curtis (BC) and Euclidean
183 distances. Statistical differences between defined groups were determined through analysis of
184 similarities (ANOSIM) and permutational multivariate analysis of variance (PERMANOVA). To
185 identify the vOTUs contributing the most to the dissimilarity between groups, similarity
186 percentages analysis (SIMPER) was employed. All diversity indices and statistical tests were
187 calculated using R (33) with the metagenome-Seq (34), Vegan (35), and ggplot2 (36) packages, as
188 well as PAST software (version 4) (37). To visualise overlap among viral OTUs detected in
189 *Triatominae* hosts, we first transformed the vOTU list (S3 Table) into a binary presence-absence
190 matrix with a custom-made R script. The binary matrix was passed to the UpSetR package (v1.4.0)
191 (38) to create an UpSet plot showing the unique and shared vOTUs across host species.

192 # Results

193 Sequencing of 23 *Triatominae* samples generated a total read length of 353.6 Gb. Following
194 quality control, the reads were *de novo* assembled into 6,748,633 contigs. Viral RNA content
195 accounted for approximately 0.0022% of the total RNA (after the removal of rRNA) per sample,
196 with an interquartile range (IQR) of 0.00021%-0.062%. Detailed sequencing and assembly metrics
197 for each sample are presented in (S4 Table). For ten Triatominae individuals, blood meal analysis
198 identified five different hosts, including wood rat, opossum, toads (likely *Incilius nebulifer*), and
199 human (S2 Table). While we cannot reliably exclude the possibility of human DNA contamination
200 during the collection and dissection process, the results unequivocally show that all analyzed
201 individuals of *T. rubida* and *H. protracta* fed on *Neotoma albigula*, whereas some *T. indictiva* and
202 *T. gerstaeckeri* utilized toads and opossums as blood meal sources.

203 # Triatominae virome composition and virus sequences annotation

204 The *de novo* assembly revealed 41 unique vOTUs, which included 40 complete or partial genomes
205 of novel viruses. Most of the Triatominae virome consists of complete or partially complete
206 genomes of single-stranded RNA (ssRNA) viruses. Among these, six negative-sense ssRNA can
207 be reliably assigned to the families *Rhabdoviridae*, *Chuviridae*, *Arenaviridae*, *Xinmoviridae,* and
208 *Orthomyxoviridae*, and the order *Elliovirales* (Fig 1). The positive-sense ssRNA viruses detected
209 belonged to the families *Virgaviridae*, *Benyviridae*, *Tombusviridae*, *Solemoviridae, Astroviridae*,
210 and *Narnaviridae*. Additionally, a single double-stranded RNA (dsRNA) virus from the family
211 *Partitiviridae* and two single-stranded DNA (ssDNA) viruses from the families *Circoviridae* and
212 *Microviridae* were detected. All five Triatominae species harbor a diverse virome. We identified
213 three to ten viral families, with *T. indictiva* and *T. sanguisuga* harboring the highest diversity (ten
214 and eight families, respectively). The viral diversity at the individual sample level includes a
215 median of four vOTUs, with an interquartile range (IQR) of four and a maximum of eleven vOTUs.

216

217 **Fig 1. Overview of the viral metagenome diversity associated with five Triatominae species**
218 **sampled from Arizona (AZ), New Mexico (NM) and Texas (TX).** The schematic phylogeny
219 represents *cytb*-derived Maximum Likelihood tree with collapsed species/population clusters (S1
220 Fig). The abbreviations in parentheses following TX refer to two nearby sampling localities in
221 Texas (detailed in S1 Table). The numbers stand for the number of samples analyzed per branch.
222 The scale bar indicates the number of identified viral families and orders per branch.

223

Negative-sense, single-stranded RNA metavirome

A total of 17 vOTUs associated with negative-sense single-stranded RNA (ssRNA-) viruses were identified across the Triatominae samples (Fig 2, S3 Table). Two vOTUs from the *Arenaviridae* (Fig 2A) family detected in *H. protracta* and *T. rubida* clustered together with Big Brushy Tank Virus (BBTV) (ACL36816.1) in the phylogeny. BBTV was previously identified in the white-throated woodrat, *Neotoma albigula*, from Arizona, suggesting ecological link between these viruses, reservoir hosts and triatomine vectors in this region. Seven vOTUs were detected within the *Chuviridae* (Fig 2B) family. One vOTU from *T. sanguisuga* comprised a complete chuvirus genome and clustered as a sister taxon to Guiyang Chuvirus 1 (UHK03046.1), previously described from the paddy ear head bug (*Cletus punctiger*) in China. Six distinct vOTUs were identified in *T. indictiva, T. sanguisuga,* and *T. gerstaeckeri*. Phylogenetic comparison showed that these vOTUs are most closely related to Sanya Chuvirus 2 (UHK03098.1), detected in a brown planthopper *Nilaparvata lugens* in China. In addition, we identified two vOTUs, corresponding to the full-length RdRp segments of an Elliovirales (Fig 2C) genome. Both vOTUs found in *T. sanguisuga* and *T. gersteackeri* shared the highest similarity with the Orius laevigatus bunyavirus 2 (PP908622.1; S3 Table). The phylogenetic analysis places them as a sister taxa to a clade of Orthohantaviruses isolated from various rodents. Within the *Orthomyxoviridae* (Fig 2D-E) family, we assembled two vOTUs corresponding to partial polymerase and nucleocapsid segments. The first, detected in both *H. protracta* and *T. indictiva* samples, with the closest relative Halyomorpha halys orthomyxo-like virus 1 (URQ09140.1). The phylogenetic analysis placed the virus within hemipteran and neuropteran orthomyxoviruses. Similar phylogenetic placement was recovered for the second vOTU found exclusively in *H. protracta*, which showed the highest similarity to Hemipteran Orthomyxo-related Virus OKIAV191 (QMP82198.1), previously isolated from hawthorn shield bug *Acanthosoma haemorrhoidale* in Germany.

249 **Fig 2. Phylogenetic trees of negative-sense, single-stranded RNA viruses detected in**
250 ***Triatominae.*** Maximum-likelihood trees were inferred from RdRp, Polymerase, and Nucleocapsid
251 protein sequences of *Xinmoviridae, Arenaviridae, Chuviridae, Rhabdoviridae,* Elliovirales, and
252 *Orthomyxoviridae.* Newly identified virus sequences in bold and support values evaluated with
253 1,000 ultrafast bootstrap replicates and 1,000 SH-aLRT tests are shown under nodes. Schematic
254 diagrams depict the genomic architecture of representative viruses, with the hatched regions
255 indicating the part of virome obtained in this study.

256

257 In the *Rhabdoviridae* (Fig 2F) family, three vOTUs were assembled from *T. indictiva, T.*
258 *sanguisuga,* and *T. rubida*. Two vOTUs corresponded to RdRp-encoding segments closely related
259 to Sanya Conocephalus Maculatus Rhabdovirus 1 (UHK03110.1), originally reported from bush
260 crickets in rice fields in China, and one vOTU to Blattodean Rhabdo-related Virus OKIAV14
261 (QMP82336.1), previously detected in red-headed cockroaches from Germany. Finally, in the
262 *Xinmoviridae* (Fig 2G) family, partial RdRp segments were assembled from *T. gerstaeckeri* and
263 *T. rubida* samples showing similarity to Hangzhou Cletus punctiger Xinmovirus 1 (UHK03158.1),
264 a virus originally characterized from rice stink bug *Cletus punctiger* in China.

265 Positive-sense, single-stranded RNA metavirome

266 A total of 15 vOTUs associated with positive-sense single-stranded RNA (ssRNA+) viruses were
267 identified (Fig 3, S3 Table). A single vOTU associated with the partial genome of the family
268 *Astroviridae* (Fig 3A) was assembled from *T. indictiva*. This vOTU showed similarity to Flumine
269 Astrovirus 3 (UQB76172.1), which was sampled from a natural water source in New Zealand.
270 However, although the phylogenetic analysis placed the vOTU within the "Flumine
271 astrovirus cluster", its position within the tree remained unclear due to low support (Fig 3A).

272

273 **Fig 3. Phylogenetic trees of positive-sense, single-stranded RNA viruses found in**
274 ***Triatominae.*** Maximum-likelihood trees were reconstructed from RdRp protein sequences of
275 *Virgaviridae, Benyviridae, Solemoviridae, Astroviridae, Tombusviridae,* and *Narnaviridae.* Node
276 support was evaluated with 1,000 ultrafast bootstrap replicates and 1,000 SH-aLRT tests, and
277 viruses identified in this study are highlighted in bold. Schematic diagrams represent the genomic
278 architecture of each virus family, with the hatched regions indicating the part of virome obtained
279 in this study.

280

281 Four vOTUs were identified in the family *Benyviridae* (Fig 3B), with two from *T. gerstaeckeri*
282 and two from *H. protracta*, all clustering with Sanya Benyvirus 1 (UHK03121.1), previously
283 isolated from *Sesamia inferens* in China. We assembled two complete *Narnaviridae* (Fig 3C)
284 genomes, from *T. indictiva*. One vOTU showed sequence identity and clustered with Serbia Narna-
285 like Virus 3 (QNS17454.1), found in *Culex pipiens* from Serbia. The other full segment showed
286 nucleotide similarity to Wuhan spider virus 7 (NC_033702.1), however the phylogenetic analysis
287 placed this vOTU with Kummerowia striata ourmiavirus (QKI29250.1) with low support values.
288 Two unrelated vOTUs corresponding to complete genomes from the family *Solemoviridae* (Fig
289 3D) were also identified. One vOTU, from *T. indictiva*, contained a fully assembled RdRp
290 sequence showing high sequence identity Atrato Sobemo-like Virus 2 (QHA33888.1), previously
291 identified in *Culex* sp. from Colombia. The second vOTU is related to both Rhodnius prolixus
292 Virus 5 and 6 (QYC92635.1, QYC92637.1), found previously in ovaries of *Rhodnius prolixus*
293 from Brazil.

294 Additionally, in both *T. sanguisuga* and *T. indictiva*, a unique vOTU related to the family
295 *Tombusviridae* (Fig 3E) was identified. This vOTU corresponded to a complete genome that

296 showed similarity to Hemipteran Tombus-related Virus (QTJ63611.1) from elongated-grass bugs,
297 however its position within the topology was not well supported. Finally, five vOTUs were
298 identified as *Virgaviridae* (Fig 3F), which include both complete and partial genomes. Three
299 vOTUs recovered from *T. gersteackeri* exhibited similarity to Xiangshan Martelli-like Virus 3
300 (UDL14011.1) and Xiangshan martelli-like virus 2 (UDL14009.1), previously identified from an
301 insect collected in China. The remaining two vOTUs were recovered from *T. sanguisuga* and *T.*
302 *indictiva and* showed similarity to the Pedersore Virga-like Virus (UYL94368.1) and Ginka Virga-
303 like Virus (UYE93845.1), which were identified in *Ochlerotatus communis* in Sweden and
304 Finland. Phylogenetically, all these vOTUs create a monophyletic cluster closely related to the
305 viruses mentioned above (Fig 3F).

306 **Single-stranded DNA metavirome**

307 Five vOTUs were identified, belonging to two families of ssDNA viruses, *Circoviridae* and
308 *Microviridae* (Fig 4, S3 Table). Four vOTUs were assigned to the *Circoviridae* (Fig 4A-B) family.
309 Three of these vOTUs detected in *T. rubida*, *T. sanguisuga,* and *T. indictiva*, showed high
310 similarity to a capsid (WGL41060.1) and replicase segments (XNO10979.1) of Circovirus sp.,
311 which were sampled from the feces of *Neophema chrysogaster* in Australia and *Ixodes ovatus* from
312 China. The additional vOTU was found in *T. rubida* and *T. indictiva*. It showed similarity to
313 Cressdnaviricota sp. (AXH74253.1), however its phylogenetic position remained unstable (Fig
314 4B). The *Microviridae* (Fig 4C) family was exclusively detected in *T. rubida* and was found to be
315 related to Microvirus sp. (WNK12593.1), sampled from the New Zealand mud snail.

316

317 **Fig 4. Phylogenetic trees of single-stranded DNA and double-stranded RNA viruses found in**
318 ***Triatominae.*** Maximum-likelihood trees were reconstructed from RdRp protein sequences of
319 *Partitiviridae,* Capsid protein of *Microviridae* and *Circoviridae,* and replication-associated protein

320 of *Circoviridae.* Identified viruses in this study are highlighted in bold letters. Node support was
321 evaluated with 1,000 ultrafast bootstrap replicates and 1,000 SH-aLRT tests and is shown in
322 branches. Schematic diagrams represent the genomic architecture of each virus family, with the
323 shaded regions indicating the sequence areas successfully recovered in this study.

324

325 Double-stranded RNA metavirome

326 Only one family of double-stranded RNA viruses, *Partitiviridae* (Fig 4D), was identified in the
327 Triatominae samples. Partial genome segments were assembled into four distinct vOTUs from the
328 species *T. sanguisuga*, *T. indictiva*, and *H. protracta* (Fig. 4, S3 Table). Three of these vOTUs,
329 exhibited high similarity to Hubei Partiti-like Virus 10 (APG78227.1) and Hubei Partiti-like Virus
330 11 (YP_009329875.1) both the viruses were previously reported in dragonflies from China. The
331 fourth vOTU, detected both in *T. sanguisuga,* and *T. indictiva,* was closely related to Hangzhou
332 partitivirus 1 (UHK03007.1), reported in dragonflies found in rice fields in China.

333 Geographic specificity of the Triatominae metavirome

334 Among the 41 vOTUs identified in this study, 26 were found exclusively in Texas, while five were
335 unique to Arizona. In Arizona, we found vOTUs belonging exclusively to the *Arenaviridae* and
336 *Microviridae* families and were detected in the species *T. rubida* and *H. protracta*. Whereas, in
337 Texas, vOTUs exclusively belonged to the order Elliovirales and the families *Chuviridae*,
338 *Tombusviridae*, *Virgaviridae*, *Astroviridae*, *Solemoviridae*, and Narna-related, which were only
339 found within the state. Notably, two vOTUs were identified in triatomines from three states,
340 corresponding to ssDNA viruses of the *Circoviridae* family. Some vOTUs showed high relative
341 abundance in various samples, particularly those related to *Solemoviridae*, and Elliovirales (S3
342 Table, S2 Fig).

Although 31 of the 41 vOTUs were geographically specific, alpha diversity indices, as measured by the Shannon diversity index, showed slight variation among specimens from the three states (mean Shannon index = 0.675, SD = 0.398), but no statistically significant differences were observed (Kruskal-Wallis test, $p > 0.05$). Similarly, richness and evenness metrics did not differ significantly between states (Kruskal-Wallis test, $p > 0.05$), indicating a broadly uniform distribution of alpha diversity indices across geographical locations. In contrast, beta diversity analyses revealed some geographic structuring. Bray-Curtis-based Non-metric multidimensional scaling (NMDS) and Euclidean-based heatmaps suggested distinct patterns in the composition and abundance of vOTUs by state. These observed groupings were statistically significant (PERMANOVA test, $p = 0.003$, $R^2 = 0.171$) (S2 Fig).

Host-specific metavirome profile in triatomines

Considering that some Triatominae species were collected exclusively in Texas and Arizona, and the observed geographic patterns might also reflect host influence in shaping the virome, we further investigated whether vOTUs shared across locations exhibited species or lineage-specific patterns. Out of the 41 vOTUs identified, 24 were detected exclusively in a single species: nine in *T. gerstaeckeri*, seven in *T. indictiva*, three each in *T. rubida* and *H. protracta*, and two in *T. sanguisuga*. The remaining 17 vOTUs were not species-specific and were shared among different species. None of these vOTUs were found in all five species. Specifically, eight vOTUs were shared between *T. indictiva* and *T. sanguisuga*; two of these vOTUs, belonging to the families *Circoviridae* and *Partitiviridae,* were shared with *T. rubida* and *H. protracta*, respectively. The other non-species-specific vOTUs were exclusively shared between pairs of species: *T. rubida* had a unique vOTU shared with *T. indictiva*, *T. sanguisuga,* and *T. gerstaeckeri*. Similarly, *T. indictiva* and *H. protracta* shared one vOTU, while *T. gerstaeckeri* and *T. sanguisuga* had two vOTUs in common (Fig 5).

368 **Fig 5. Alpha and beta diversity of vOTUs in Triatominae natural populations from southern**
369 **United States.** A) NMDS based on Bray-Curtis distances, showing clustering patterns by
370 Triatominae species (Stress = 0.097, PERMANOVA, $p < 0.001$; $R^2 = 0.502$). B) Shannon diversity
371 index of vOTUs by Triatominae species (Kruskal-Wallis test, $p = 0.399$). C) Venn diagram
372 displaying the number of shared vOTUs among Triatominae species. D) Heatmap based on
373 normalized read counts of vOTUs and their Euclidean-based clustering among Triatominae
374 species.

375

376 The Shannon diversity index of vOTUs revealed no significant differences among Triatominae
377 species (Kruskal-Wallis test, $p > 0.05$). In contrast, significant differences were observed in vOTUs
378 richness and evenness across the five Triatominae species (Kruskal-Wallis test, $p < 0.05$). Beta
379 diversity analyses revealed distinct clustering patterns in vOTUs composition and abundance
380 among species, with statistically significant groupings (PERMANOVA test, $p = 0.001$, $R^2 = 0.502$;
381 ANOSIM test, $p = 0.0001$, R = 0.713) (Fig 5). Notably, samples from different Triatominae species
382 generally clustered distinctly, and the virome of *T. indictiva* exhibited the highest diversity of
383 vOTUs and clustered more distant from the virome of other species, aligning more closely with *T.*
384 *sanguisuga*, as these two species shared the most vOTUs.

385 The SIMPER analysis showed that a group of six vOTUs explained more than 74% of the
386 dissimilarity in the virome composition among the species, highlighting the contribution of the
387 species-specific vOTUs. Solemo_vOTU25, unique to *T. indictiva*, was the main contributor,
388 explaining 26.8% of the dissimilarity. It was followed by Solemo_vOTU26 (14.7%), found
389 exclusively in *T. sanguisuga*. Additionally, Rhabdo_vOTU41 was present only in *T. rubida*, while
390 Arena_vOTU15 was unique to *H. protracta*. Narna_vOTU24, contributing 6.3% to the
391 dissimilarity, was also found only in *T. indictiva*, and Beny_vOTU19, with a contribution of 3%,
392 was unique to *T. gerstaeckeri*. Furthermore, the Ellio_vOTU13 and Ellio_vOTU14 contributed an

393 additional 11.37% and 7.89%, respectively, and were more abundant in *T. sanguisuga* and *T.*
394 *gerstaeckeri* (S5 Table).

395 Sex, tissue, and T. cruzi infection status determinants

396 A uniform distribution of alpha viral diversity indices was observed across tissue type, sex, and *T.*
397 *cruzi* infection status. Shannon, richness, and evenness indices did not show statistically significant
398 differences among these factors (Kruskal-Wallis tests, $p > 0.05$). L5 specimens exhibited a higher
399 Shannon index (Mean Shannon index = 1.18, SD = 0.163) compared to males (Mean Shannon
400 index = 0.548, SD = 0.408). The Beta diversity analyses revealed no distinct clustering patterns in
401 vOTU composition or abundance based on sex, tissue type, or *T. cruzi* infection status
402 (PERMANOVA test, $p > 0.05$; ANOSIM test, $p > 0.05$) (S1 Fig).

403 Remarkably, 24 of the vOTUs identified in this study were found in reproductive tissue.
404 Additionally, two vOTUs from *T. cruzi*-positive samples, which belong to the viral families
405 *Orthomyxoviridae* and *Chuviridae*, were not exclusive to the infected individuals.

406 Discussion

407 Our study advances the field by characterizing the virome of five Triatominae species from natural
408 populations in the southern United States: *Triatoma rubida, Hospesneotomae protracta, Triatoma*
409 *indictiva, Triatoma sanguisuga*, and *Triatoma gerstaeckeri*. We identified a diverse range of viral
410 metagenomes, including negative- and positive-sense ssRNA, dsRNA, and small circular ssDNA
411 viruses from various families. Our findings indicate that viral composition and abundance are
412 primarily shaped by host species identity, suggesting a species-specific viral metagenome profile.

413 A significant proportion of the identified sequences belonged to *Chuviridae*, *Virgaviridae*,
414 *Benyviridae*, *Partitiviridae,* and *Rhabdoviridae.* While some members of these families infect non-

415 arthropod hosts, the phylogenetic analyses showed that most viral sequences clustered within
416 clades of previously described arthropod-associated viruses, reinforcing their specificity to insects.
417 Additionally, these viral taxa have been consistently reported in insect viromes (2, 39, 40) and, in
418 some cases, as endogenous viral elements (41-44), further supporting their specificity to
419 arthropods.

420 Interestingly, viral sequences from families primarily associated with plant infections were also
421 detected. Solemoviruses, in particular, have been repeatedly identified in triatomines. Rhodnius
422 prolixus virus 5 (RpV5) and Rhodnius prolixus virus 6 (RpV6), initially reported in *R. prolixus*
423 (14), have since been found in *Rhodnius robustus* in French Guiana and *Rhodnius neglectus* in
424 Brazil (15). This study expands the known host range of *Solemoviridae* by identifying Triatoma
425 indictiva solemoviridae vOTU25 and Triatoma sanguisuga solemoviridae vOTU26. The recurring
426 presence of these viruses in multiple triatomine species across different regions suggests a
427 widespread distribution throughout the Americas. Additionally, we identified plant-virus-related
428 sequences, including Triatoma spp. tombusviridae vOTU22 and Benyviruses (Beny_vOTU18-21)
429 in *T. gerstaeckeri* and *H. protracta*. While *Benyviridae* members are primarily associated with
430 plant infections (45), their detection in insects has been documented (46-48), though their role in
431 non-plant hosts remains unclear (49). While our study does not assess whether triatomines
432 contribute to plant virus transmission, our findings highlight a potential ecological role warranting
433 further investigation.

434 Vertebrate-associated viruses within the family *Arenaviridae* were also identified, specifically
435 Hospesneotomae protracta arenaviridae vOTU15 and Triatoma rubida arenaviridae vOTU16.
436 These viruses are phylogenetically related to BBTV, previously isolated from *Neotoma albigula*
437 in Arizona (50, 51). BBTV belongs to the genus *Mammarenavirus*, which includes 51 species
438 primarily associated with rodents (52). Notably, these arenaviruses were detected exclusively in
439 *T. rubida* and *H. protracta* from Arizona. Consistent with these findings, we detected *Neotoma*
440 blood in all analyzed individuals of *T. rubida* and *H. protracta*, further supporting the association

441 between arenavirus sequences and rodent-derived blood meals. The presence of *Arenaviridae*
442 vOTUs in the intestinal tract suggests that triatomines may act as mechanical carriers, dispersing
443 viral particles through feces when interacting with rodent hosts. However, it remains unclear
444 whether triatomines contribute to the natural transmission cycles of *Mammarenavirus*, or whether
445 these findings simply reflect transient viral passage. Further investigation is needed to elucidate
446 the role of triatomines in *Mammarenavirus* epidemiology.

447 In addition to RNA virome, we identified small circular DNA viruses from *Circoviridae* and
448 *Microviridae* family. Circovirus sequences were phylogenetically related to viruses previously
449 found in environmental samples and insects. Although most known *Circoviridae* members infect
450 vertebrates (53) *Cyclovirus* species have also been identified in arthropods (54). Their presence in
451 hematophagous insects may be linked to the ingestion of viremic blood (55). Additionally,
452 circoviruses have been detected circulating in hosts from the study region (56, 57), which could
453 explain their presence in triatomines. In contrast, *Microviridae* are bacteriophages, likely
454 representing components of the bacterial microbiome within triatomines. These bacteriophages
455 have been consistently detected in insect microbiomes (58, 59), including species from regions
456 examined in this study (60).

457 Analysis of viral diversity and abundance revealed significant differences in vOTU richness and
458 evenness among the five *Triatominae* species. Beta diversity analyses consistently showed
459 species-specific clustering patterns, while sex and tissue type had minimal influence. These
460 findings align with previous studies on mosquitoes and ticks, where virome richness and diversity
461 primarily vary at the species level (3, 61-66). Beyond species identity, host interactions may also
462 play a role in shaping virome composition in *Triatominae*, as observed in other hematophagous
463 arthropods such as mosquitoes and ticks, where host associations influence virome structure (3,
464 64). Here, the detection of viral sequences from the families *Solemoviridae* and *Arenaviridae*
465 suggests that host interactions contribute to virome composition in Triatominae. Notably, Triatoma

466 indictiva solemoviridae vOTU25 and Triatoma sanguisuga solemoviridae vOTU26 were identified
467 as key drivers of differences in viral abundance and composition among the *Triatominae* species.

468 Another factor known to influence insect virome composition is geography (5, 64, 67). Our
469 analysis revealed that although 31 of the 41 vOTUs were geographically specific, beta diversity
470 analyses identified significant geographic structuring among specimens from Arizona, Texas, and
471 New Mexico. Nevertheless, since some *Triatominae* species were collected exclusively in Texas
472 and Arizona, the observed geographic patterns may, in part, reflect the influence of host species
473 identity on virome composition. These findings suggest that while overall viral richness and
474 evenness remain relatively stable across regions, differences in virome composition may primarily
475 be shaped by species identity, as mentioned earlier, without ruling out the complex interplay of
476 localized environmental factors, host distributions, and ecological interactions. Future research
477 should aim to disentangle the relative contributions of these factors, refining our understanding of
478 virus-host dynamics in *Triatominae*.

479 Finally, a key challenge in viral metagenome studies is differentiating actively replicating viruses
480 from endogenous viral elements (EVEs) (68). EVEs arise from viral insertions into host genomes,
481 complicating their distinction from exogenous viruses (69). Although complete RNA virus
482 genomes can occasionally integrate into a host genome, such events are relatively rare. EVEs are
483 typically fragmented viral sequences with disrupted ORFs (70). In our study, however, most of the
484 viral sequences we identified are full-length or nearly full-length coding regions of RNA viral
485 genomes. This suggests that these sequences are unlikely to represent EVEs expressed by the host.
486 Nonetheless, addressing this issue requires complementary approaches, such as whole-genome
487 sequencing (WGS) and small RNA sequencing, to assess viral replication and host antiviral
488 responses (63, 71). Incorporating these methodologies into future research will enhance our ability
489 to distinguish EVEs from actively replicating viruses in Triatominae, thereby contributing to a
490 deeper understanding of virus-triatomine dynamics.

Conclusion

This study offers the most comprehensive characterization of *Triatominae* viromes to date, significantly advancing our understanding of viral diversity, host specificity, and potential ecological interactions. Our findings reveal species-specific virome structures and underscore the critical role of host identity in shaping viral composition. The detection of viruses associated with insects, plants, and vertebrates highlights the intricate nature of viral interactions in these hematophagous hosts. Given the complexity of these interactions, further research is essential to explore the functional implications of these viral components and their potential impact on Triatominae ecology, as well as their effects on vector competence.

Acknowledgements

We would like to acknowledge our collaborators that contributed to the sampling efforts Robert L. Smith, Walter Roachell and Alvaro Romero. The work was supported by the Czech Science Foundation (grant number 21-10185M).

Conflict of interest

The authors declare that there was no potential conflict of interest.

Author Contributions

Conceptualization and supervision: EN. Material preparation: JCH and JM. Data analysis: IH and JCH. Writing – original draft: IH, JCH, and EN. Writing – review and editing: IH, JCH, JM and EN

Data Availability

All relevant data are within the manuscript and its Supporting Information files. Raw RNA-seq reads have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1279500. Viral genome assemblies have been deposited in GenBank under accession numbers PX139058 - PX139102.

References

1. Bolling BG, Weaver SC, Tesh RB, Vasilakis N. Insect-specific virus discovery: significance for the arbovirus community. Viruses. 2015;7(9):4911-28.

2. de Almeida JP, Aguiar ER, Armache JN, Olmo RP, Marques JT. The virome of vector mosquitoes. Current opinion in virology. 2021;49:7-12.

3. Ni X-B, Cui X-M, Liu J-Y, Ye R-Z, Wu Y-Q, Jiang J-F, et al. Metavirome of 31 tick species provides a compendium of 1,801 RNA virus genomes. Nature microbiology. 2023;8(1):162-73.

4. Vandegrift KJ, Kapoor A. The ecology of new constituents of the tick virome and their relevance to public health. Viruses. 2019;11(6):529.

5. Pan Y-F, Zhao H, Gou Q-Y, Shi P-B, Tian J-H, Feng Y, et al. Metagenomic analysis of individual mosquito viromes reveals the geographical patterns and drivers of viral diversity. Nature Ecology & Evolution. 2024;8(5):947-59.

6. Cardoso MA, Brito TFd, Brito IAdA, Berni MA, Coelho VL, Pane A. The neglected virome of triatomine insects. Frontiers in Tropical Diseases. 2022;3:828712.

7. Muscio O, LaTorre J, Scodeller E. Small nonoccluded viruses from triatomine bug *Triatoma infestans* (Hemiptera: Reduviidae). Journal of invertebrate pathology. 1987;49(2):218-20.

8. Muscio O, La Torre J, Scodeller E. Characterization of Triatoma virus, a picorna-like virus isolated from the triatomine bug *Triatoma infestans*. Journal of general virology. 1988;69(11):2929-34.

535 9. Rozas-Dennis G, Cazzaniga N. Effects of Triatoma virus (TrV) on the fecundity and
536 moulting of *Triatoma infestans* (Hemiptera: Reduviidae). Annals of Tropical Medicine &
537 Parasitology. 2000;94(6):633-41.

538 10. Susevich ML, Marti GA, Serena MS, Echeverría MG. New Triatoma virus hosts in wild
539 habitats of Argentina. Journal of invertebrate pathology. 2012;110(3):405-7.

540 11. Marti GA, Echeverria MG, Susevich ML, Becnel JJ, Pelizza SA, García JJ. Prevalence
541 and distribution of parasites and pathogens of Triatominae from Argentina, with emphasis on
542 *Triatoma infestans* and Triatoma virus TrV. Journal of invertebrate pathology. 2009;102(3):233-
543 7.

544 12. Marti GA, Ragone P, Balsalobre A, Ceccarelli S, Susevich ML, Diosque P, et al. Can
545 Triatoma virus inhibit infection of *Trypanosoma cruzi* (Chagas, 1909) in Triatoma infestans (Klug)?
546 A cross infection and co-infection study. Journal of invertebrate pathology. 2017;150:101-5.

547 13. Muscio OA, La Torre JL, Bonder MA, Scodeller EA. Triatoma virus pathogenicity in
548 laboratory colonies of *Triatoma infestans* (Hemiptera: Reduviidae). Journal of medical
549 entomology. 1997;34(3):253-6.

550 14. Brito TFd, Coelho VL, Cardoso MA, Brito IAdA, Berni MA, Zenk FL, et al. Transovarial
551 transmission of a core virome in the Chagas disease vector *Rhodnius prolixus*. PLoS Pathogens.
552 2021;17(8):e1009780.

553 15. Bourke BP, de Oliveira J, Ergunay K, Linton Y-M. Leveraging transcriptome sequence
554 read archives for virus detection in wild and colony populations of triatomines (Hemiptera:
555 Reduviidae: Triatominae). Archives of Virology. 2024;169(10):215.

556 16. Bolger A, Giorgi F. Trimmomatic: A flexible read trimming tool for Illumina NGS data.
557 Bioinformatics. 2014;30(15):2114-20.

558 17. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile
559 metagenomic assembler. Genome research. 2017;27(5):824-34.

560 18. Bigot T, Temmam S, Pérot P, Eloit M. RVDB-prot, a reference viral protein database and
561 its HMM profiles. F1000Research 8: 1–13. 2020.

562 19. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND.
563 Nature methods. 2015;12(1):59-60.

564 20. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein
565 families database in 2019. Nucleic acids research. 2019;47(D1):D427-D32.

566 21. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer EL, et al.
567 Pfam: The protein families database in 2021. Nucleic acids research. 2021;49(D1):D412-D9.

568 22. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity
569 searching. Nucleic acids research. 2011;39(suppl_2):W29-W37.

570 23. Wernersson R. Virtual Ribosome—a comprehensive DNA translation tool with support for
571 integration of sequence feature annotation. Nucleic acids research. 2006;34(suppl_2):W385-W8.

572 24. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool.
573 Journal of molecular biology. 1990;215(3):403-10.

574 25. Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB, et al. Virus
575 taxonomy in the age of metagenomics. Nature Reviews Microbiology. 2017;15(3):161-8.

576 26. Alvarez-Jarreta J, Amos B, Aurrecoechea C, Bah S, Barba M, Barreto A, et al.
577 VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center in 2023.
578 Nucleic acids research. 2023:gkad1003-gkad.

579 27. Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology
580 detection and structure prediction. Nucleic acids research. 2005;33(suppl_2):W244-W8.

581 28. Kans J. Entrez direct: e-utilities on the UNIX command line. Entrez programming utilities
582 help [Internet]: National Center for Biotechnology Information (US); 2023.

583 29. Katoh K, Misawa K, Kuma Ki, Miyata T. MAFFT: a novel method for rapid multiple
584 sequence alignment based on fast Fourier transform. Nucleic acids research. 2002;30(14):3059-
585 66.

586 30. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment
587 trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972-3.

588 31. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et
589 al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era.
590 Molecular biology and evolution. 2020;37(5):1530-4.

591 32. Roux S, Adriaenssens EM, Dutilh BE, Koonin EV, Kropinski AM, Krupovic M, et al.
592 Minimum information about an uncultivated virus genome (MIUViG). Nature biotechnology.
593 2019;37(1):29-37.

594 33. Team RC. R: A Language and Environment for Statistical Computing. 4.1.2 ed2021.

595 34. Paulson JN, Pop M, Bravo HC. metagenomeSeq: Statistical analysis for sparse high-
596 throughput sequencing. Bioconductor package. 2013;1(0):191.

597 35. Oksanen J-p. vegan: Community Ecology Package-R package version 1.17-8.
598 http://CRAN R-project org/package= vegan. 2011.

599 36. Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, et al. ggplot2: create
600 elegant data visualisations using the grammar of graphics. 2020. URL https://CRAN R-project
601 org/package= ggplot2 R package version. 2022;3(0):147.

602 37. Harper D. PAST: paleontological statistics software package for education and data
603 analysis. Palaeontol Electron. 2001;4:4.

604 38. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of
605 intersecting sets and their properties. Bioinformatics. 2017;33(18):2938-40.

606 39. Hernandez-Valencia JC, Muñoz-Laiton P, Gómez GF, Correa MM. A systematic review
607 on the viruses of Anopheles mosquitoes: the potential importance for public health. Tropical
608 Medicine and Infectious Disease. 2023;8(10):459.

609 40. Moming A, Bai Y, Wang J, Zhang Y, Tang S, Fan Z, et al. The known and unknown of
610 global tick-borne viruses. Viruses. 2024;16(12):1807.

611 41. Gilbert C, Belliardo C. The diversity of endogenous viral elements in insects. Current
612 opinion in insect science. 2022;49:48-55.

613 42. Hernandez-Valencia JC, Munoz-Laiton P, Gomez GF, Correa MM. Evidence of
614 endogenous non-retroviral RNA virus sequences into the genome and transcriptome of the
615 malaria vector *Anopheles darlingi*. Acta Tropica. 2024;260:107469.

616 43. Russo AG, Kelly AG, Enosi Tuipulotu D, Tanaka MM, White PA. Novel insights into
617 endogenous RNA viral elements in Ixodes scapularis and other arbovirus vector genomes. Virus
618 Evolution. 2019;5(1):vez010.

619 44. Wallau GL. RNA virus EVEs in insect genomes. Current Opinion in Insect Science.
620 2022;49:42-7.

621 45. Gilmer D, Ratti C, Consortium IR. ICTV virus taxonomy profile: *Benyviridae*. Journal of
622 General Virology. 2017;98(7):1571-2.

623 46. Debat H, Farrher ES, Bejerman N. Insights into the RNA virome of the corn leafhopper
624 *Dalbulus maidis*, a major emergent threat of Maize in Latin America. Viruses. 2024;16(10):1583.

625 47. Wu H, Pang R, Cheng T, Xue L, Zeng H, Lei T, et al. Abundant and diverse RNA viruses
626 in insects revealed by RNA-Seq analysis: ecological and evolutionary implications. Msystems.
627 2020;5(4):10.1128/msystems. 00039-20.

628 48. Huang H-J, Ye Z-X, Wang X, Yan X-T, Zhang Y, He Y-J, et al. Diversity and infectivity of
629 the RNA virome among different cryptic species of an agriculturally important insect vector:
630 whitefly *Bemisia tabaci*. npj Biofilms and Microbiomes. 2021;7(1):43.

631 49. Solovyev AG, Morozov SY. Uncovering plant virus species forming novel provisional
632 taxonomic units related to the Family *Benyviridae*. Viruses. 2022;14(12):2680.

633 50. Cajimat MN, Milazzo ML, Haynie ML, Hanson JD, Bradley RD, Fulhorst CF. Diversity and
634 phylogenetic relationships among the North American Tacaribe serocomplex viruses (family
635 *Arenaviridae*). Virology. 2011;421(2):87-95.

636 51. Milazzo ML, Cajimat MN, Haynie ML, Abbott KD, Bradley RD, Fulhorst CF. Diversity
637 among Tacaribe serocomplex viruses (family *Arenaviridae*) naturally associated with the white-
638 throated woodrat (*Neotoma albigula*) in the southwestern United States. Vector-borne and
639 zoonotic diseases. 2008;8(4):523-40.

640 52. Radoshitzky SR, Buchmeier MJ, Charrel RN, Gonzalez J-PJ, Günther S, Hepojoki J, et al.
641 ICTV virus taxonomy profile: *Arenaviridae* 2023. Journal of General Virology.
642 2023;104(9):001891.

643 53. Breitbart M, Delwart E, Rosario K, Segalés J, Varsani A, Consortium IR. ICTV virus
644 taxonomy profile: *Circoviridae*. Journal of General Virology. 2017;98(8):1997-8.

645 54. Garigliany M-M, Börstler J, Jöst H, Badusche M, Desmecht D, Schmidt-Chanasit J, et al.
646 Characterization of a novel circo-like virus in *Aedes vexans* mosquitoes from Germany: evidence
647 for a new genus within the family *Circoviridae*. Journal of General Virology. 2015;96(4):915-20.

648 55. Ha Z, Li J-F, Xie C-Z, Li C-H, Zhou H-N, Zhang Y, et al. First detection and genomic
649 characterization of porcine circovirus 3 in mosquitoes from pig farms in China. Veterinary
650 microbiology. 2020;240:108522.

651 56. Cerna GM, Serieys LE, Riley SP, Richet C, Kraberger S, Varsani A. A circovirus and
652 cycloviruses identified in feces of bobcats (*Lynx rufus*) in California. Archives of virology.
653 2023;168(1):23.

654 57. Hess SC, Weiss KC, Custer JM, Lewis JS, Kraberger S, Varsani A. Identification of small
655 circular DNA viruses in coyote fecal samples from Arizona (USA). Archives of virology.
656 2024;169(1):12.

657 58. Flynn PJ, Moreau CS. Viral diversity and co-evolutionary dynamics across the ant
658 phylogeny. Molecular Ecology. 2024;33(19):e17519.

659 59. Kraberger S, Cook CN, Schmidlin K, Fontenele RS, Bautista J, Smith B, et al. Diverse
660 single-stranded DNA viruses associated with honey bees (*Apis mellifera*). Infection, Genetics and
661 Evolution. 2019;71:179-88.

662 60. Bandoo RA, Kraberger S, Ozturk C, Lund MC, Zhu Q, Cook C, et al. Identification of
663 Diverse Bacteriophages Associated with Bees and Hoverflies. Viruses. 2025;17(2):201.

664 61. Calle-Tobón A, Pérez-Pérez J, Forero-Pineda N, Chávez OT, Rojas-Montoya W, Rúa-
665 Uribe G, et al. Local-scale virome depiction in Medellín, Colombia, supports significant differences
666 between *Aedes aegypti* and *Aedes albopictus*. PLoS One. 2022;17(7):e0263143.

667 62. Feng Y, Gou Q-y, Yang W-h, Wu W-c, Wang J, Holmes EC, et al. A time-series meta-
668 transcriptomic analysis reveals the seasonal, host, and gender structure of mosquito viromes.
669 Virus Evolution. 2022;8(1):veac006.

670 63. Li C, Liu S, Zhou H, Zhu W, Cui M, Li J, et al. Metatranscriptomic sequencing reveals host
671 species as an important factor shaping the mosquito virome. Microbiology Spectrum.
672 2023;11(2):e04655-22.

673 64. Liu Q, Cui F, Liu X, Fu Y, Fang W, Kang X, et al. Association of virome dynamics with
674 mosquito species and environmental factors. Microbiome. 2023;11(1):101.

675 65. Tian D, Ye R-Z, Li Y-Y, Wang N, Gao W-Y, Wang B-H, et al. Virome specific to tick genus
676 with distinct ecogeographical distribution. Microbiome. 2025;13(1):57.

677 66. Shi C, Beller L, Deboutte W, Yinda KC, Delang L, Vega-Rúa A, et al. Stable distinct core
678 eukaryotic viromes in different mosquito species from Guadeloupe, using single mosquito viral
679 metagenomics. Microbiome. 2019;7(1):121.

680 67. Hernandez-Valencia JC, Gómez GF, Correa MM. Metagenomic analysis evidences a core
681 virome in *Anopheles darlingi* from three contrasting Colombian ecoregions. PLoS One.
682 2025;20(4):e0320593.

683 68. Ritsch M, Brait N, Harvey E, Marz M, Lequime S. Endogenous viral elements: insights into
684 data availability and accessibility. Virus Evolution. 2024;10(1):veae099.

685 69. Brait N, Hackl T, Morel C, Exbrayat A, Gutierrez S, Lequime S. A tale of caution: How
686 endogenous viral elements affect virus discovery in transcriptomic data. Virus evolution.
687 2024;10(1):vead088.

688 70. Katzourakis A, Gifford RJ. Endogenous viral elements in animal genomes. PLoS genetics.
689 2010;6(11):e1001191.

690 71. Santos D, Feng M, Kolliopoulou A, Taning CN, Sun J, Swevers L. What are the functional
691 roles of piwi proteins and piRNAs in insects? Insects. 2023;14(2):187.

692 # Supporting information

693 **S1 Table. Sample metadata and T. cruzi meta-assembly screening**. Summary of 23
694 Triatominae individuals with species, sex, geographic location, tissue type, QC, and *T. cruzi*
695 detection

696 **S2 Table. Blood meal analysis of Triatominae gut tissues.** Twelve DNA isolates from gut
697 samples analysed using vertebrate 12S rRNA gene sequences to identify blood meal sources.

698 **S3 Table. Viral Operational Taxonomic Units (vOTUs)**. List of vOTUs detected in Triatominae
699 natural populations from the southern United States.

700 **S4 Table. Sequencing metrics.** Summary of raw sequencing reads, reads retained after quality
701 control, total number of assembled contigs, and number of viral reads per sample.

702 **S5 Table. SIMPER analysis**. Contribution of each Viral Operational Taxonomic Units (vOTUs)
703 to the Bray-Curtis dissimilarities observed in Triatominae natural populations between species (*H.*
704 *protracta, T. rubida, T. indictiva, T. sanguisuga, T. gersteackeri*).

705 **S1 Fig. Host phylogeny of sampled *Triatominae* species.** Maximum-likelihood tree based on
706 COI gene from the five triatomine species included in this study (*Triatoma rubida,*
707 *Hospesneotomae protracta, Triatoma gerstaeckeri, Triatoma sanguisuga,* and *Triatoma*
708 *indictiva*). Multiple samples from the same species are shown where available and Node labels
709 indicate bootstrap support values (100 replicates).

710 **S2 Fig. Alpha and beta diversity of vOTUs in natural populations of Triatoma.** A) Non-metric
711 multidimensional scaling (NMDS) based on Bray-Curtis distances, showing clustering patterns by
712 study location (PERMANOVA, $p = 0.003$; $R^2 = 0.171$), sex (PERMANOVA, $p = 0.426$; $R^2 =$
713 0.093), and tissue (PERMANOVA, $p = 0.97$; $R^2 = 0.138$). B) Shannon diversity index of vOTUs
714 by location (Kruskal-Wallis test, $p = 0.816$), sex (Kruskal-Wallis test, $p = 0.320$), and tissue
715 (Kruskal-Wallis test, $p = 0.744$). C) Venn diagrams displaying the number of shared vOTUs among
716 triatomine populations, grouped by location, sex, and tissue type.

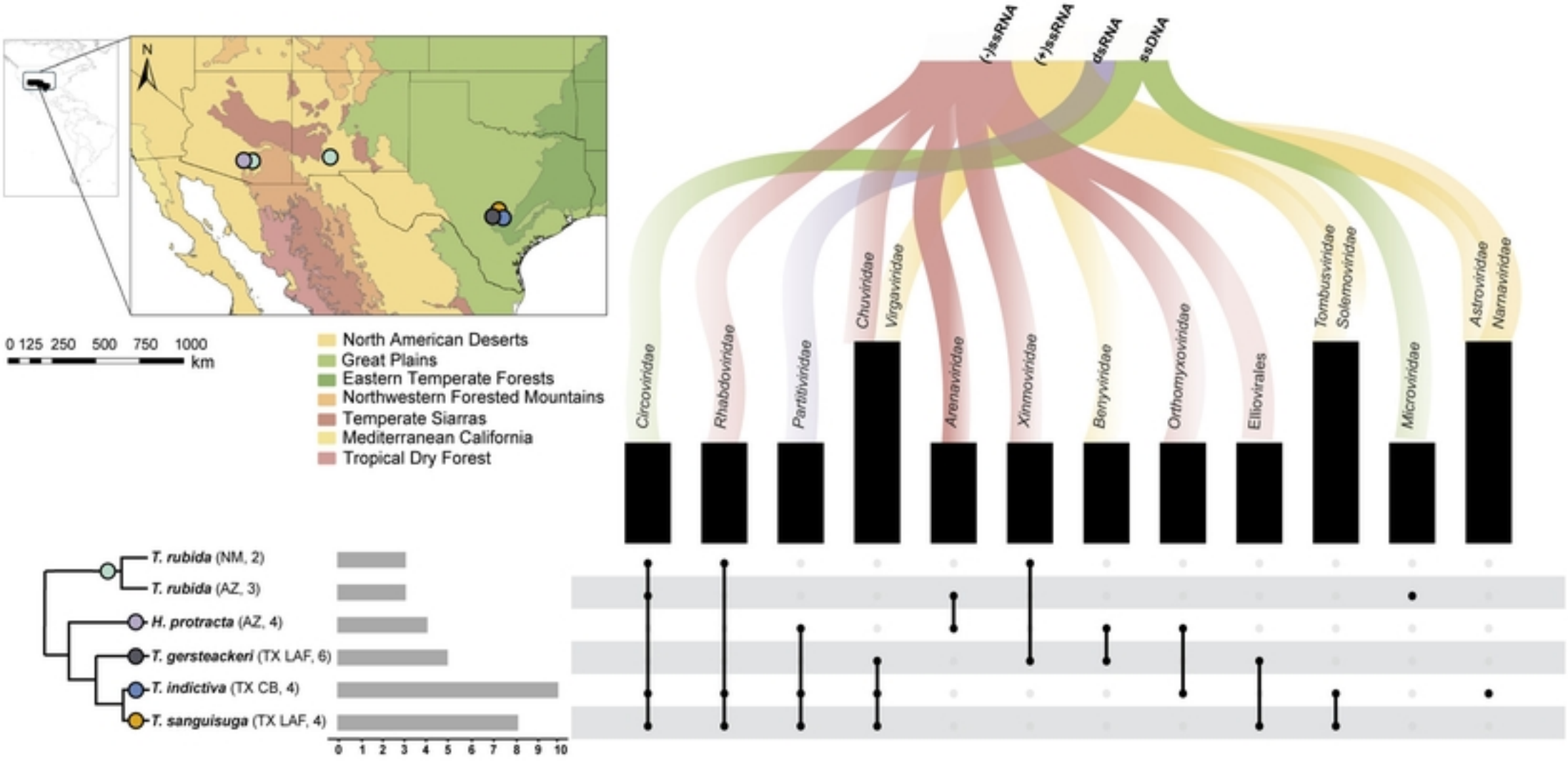
N
0 125 250 500 750 1000
km
North American Deserts
Great Plains
Eastern Temperate Forests
Northwestern Forested Mountains
Temperate Siarras
Mediterranean California
Tropical Dry Forest
(-)ssRNA
(+)ssRNA
dsRNA
ssDNA
Circoviridae
Rhabdoviridae
Partitiviridae
Chuviridae
Virgaviridae
Arenaviridae
Xinmoviridae
Benyviridae
Orthomyxoviridae
Elliovirales
Tombusviridae
Solemoviridae
Microviridae
Astroviridae
Narnaviridae
T. rubida (NM, 2)
T. rubida (AZ, 3)
H. protracta (AZ, 4)
T. gersteackeri (TX LAF, 6)
T. indictiva (TX CB, 4)
T. sanguisuga (TX LAF, 4)
0 1 2 3 4 5 6 7 8 9 10

Figure 1

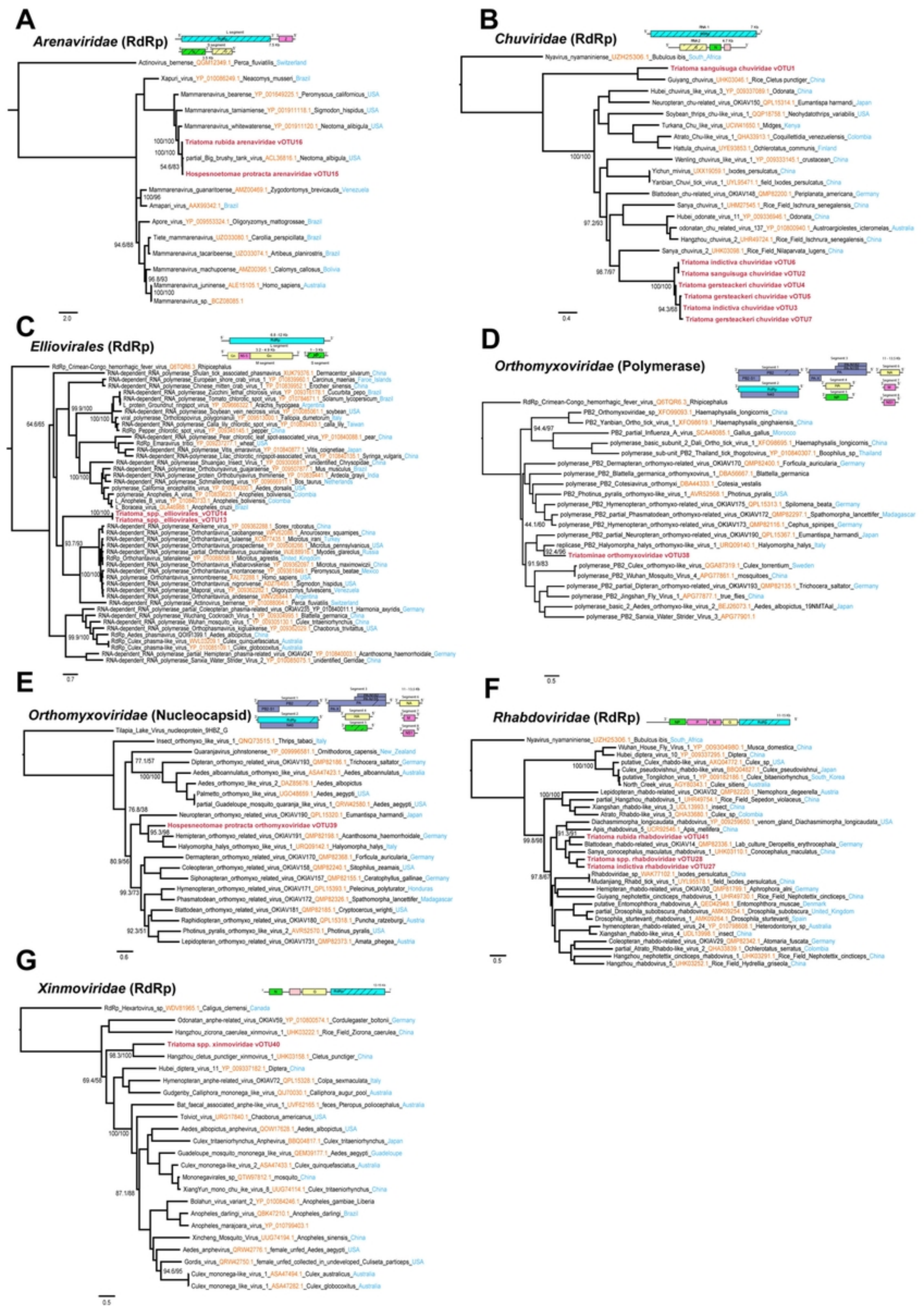
A
Arenaviridae (RdRp)
Triatoma rubida arenaviridae vOTU16
Hospesnoetomae protracta arenaviridae vOTU15
B
Chuviridae (RdRp)
Triatoma sanguisuga chuviridae vOTU1
Triatoma indictiva chuviridae vOTU6
Triatoma sanguisuga chuviridae vOTU2
Triatoma gersteackeri chuviridae vOTU4
Triatoma gersteackeri chuviridae vOTU5
Triatoma indictiva chuviridae vOTU3
Triatoma gersteackeri chuviridae vOTU7
C
Elliovirales (RdRp)
Triatoma_spp._elliovirales_vOTU14
Triatoma_spp._elliovirales_vOTU13
D
Orthomyxoviridae (Polymerase)
Triatominae orthomyxoviridae vOTU38
E
Orthomyxoviridae (Nucleocapsid)
Hospesneotomae protracta orthomyxoviridae vOTU39
F
Rhabdoviridae (RdRp)
Triatoma rubida rhabdoviridae vOTU41
Triatoma spp. rhabdoviridae vOTU28
Triatoma indictiva rhabdoviridae vOTU27
G
Xinmoviridae (RdRp)
Triatoma spp. xinmoviridae vOTU40

Figure 2

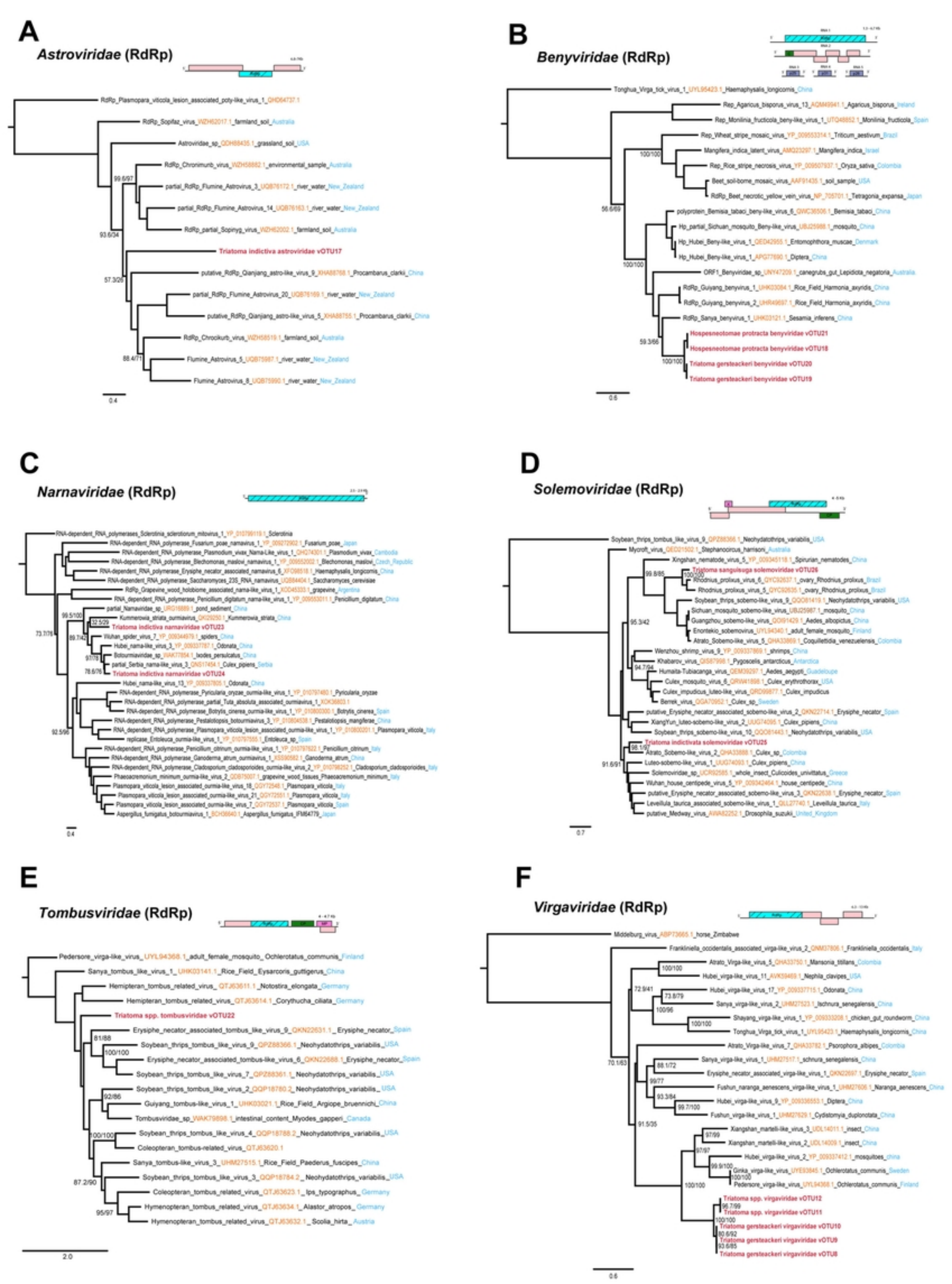
A
Astroviridae (RdRp)
Triatoma indictiva astroviridae vOTU17
B
Benyviridae (RdRp)
Hospesneotomae protracta benyviridae vOTU21
Hospesneotomae protracta benyviridae vOTU18
Triatoma gersteackeri benyviridae vOTU20
Triatoma gersteackeri benyviridae vOTU19
C
Narnaviridae (RdRp)
Triatoma indictiva narnaviridae vOTU23
Triatoma indictiva narnaviridae vOTU24
D
Solemoviridae (RdRp)
Triatoma sanguisuga solemoviridae vOTU26
Triatoma indictivata solemoviridae vOTU25
E
Tombusviridae (RdRp)
Triatoma spp. tombusviridae vOTU22
F
Virgaviridae (RdRp)
Triatoma spp. virgaviridae vOTU12
Triatoma spp. virgaviridae vOTU11
Triatoma gersteackeri virgaviridae vOTU10
Triatoma gersteackeri virgaviridae vOTU9
Triatoma gersteackeri virgaviridae vOTU8

Figure 3

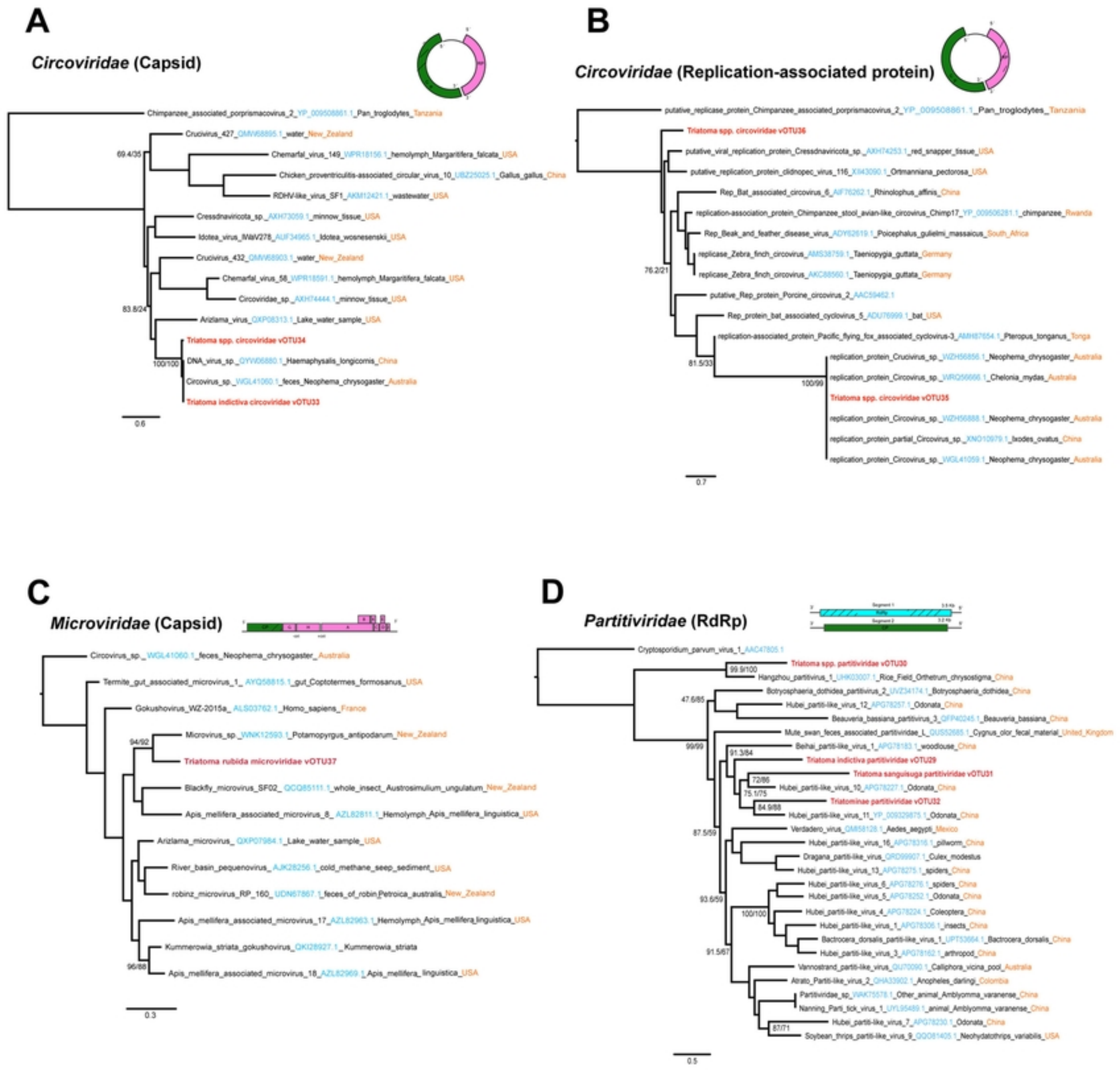
A
Circoviridae (Capsid)
Chimpanzee_associated_porprismacovirus_2_YP_009508861.1_Pan_troglodytes_Tanzania
Crucivirus_427_QMW68895.1_water_New_Zealand
69.4/35
Chemarfal_virus_149_WPR18156.1_hemolymph_Margaritifera_falcata_USA
Chicken_proventriculitis-associated_circular_virus_10_UBZ25025.1_Gallus_gallus_China
RDHV-like_virus_SF1_AKM12421.1_wastewater_USA
Cressdnaviricota_sp._AXH73059.1_minnow_tissue_USA
Idotea_virus_IWaV278_AUF34965.1_Idotea_wosnesenskii_USA
Crucivirus_432_QMW68903.1_water_New_Zealand
Chemarfal_virus_58_WPR18591.1_hemolymph_Margaritifera_falcata_USA
Circoviridae_sp._AXH74444.1_minnow_tissue_USA
83.8/24
Ariziama_virus_QXP08313.1_Lake_water_sample_USA
Triatoma spp. circoviridae vOTU34
100/100
DNA_virus_sp._QYW06880.1_Haemaphysalis_longicornis_China
Circovirus_sp._WGL41060.1_feces_Neophema_chrysogaster_Australia
Triatoma indictiva circoviridae vOTU33
0.6
B
Circoviridae (Replication-associated protein)
putative_replicase_protein_Chimpanzee_associated_porprismacovirus_2_YP_009508861.1_Pan_troglodytes_Tanzania
Triatoma spp. circoviridae vOTU36
putative_viral_replication_protein_Cressdnaviricota_sp._AXH74253.1_red_snapper_tissue_USA
putative_replication_protein_clidnopec_virus_116_XI43090.1_Ortmanniana_pectorosa_USA
Rep_Bat_associated_circovirus_6_AIF76262.1_Rhinolophus_affinis_China
replication-association_protein_Chimpanzee_stool_avian-like_circovirus_Chimp17_YP_009506281.1_chimpanzee_Rwanda
Rep_Beak_and_feather_disease_virus_ADY82619.1_Poicephalus_gulielmi_massaicus_South_Africa
replicase_Zebra_finch_circovirus_AMS38759.1_Taeniopygia_guttata_Germany
76.2/21
replicase_Zebra_finch_circovirus_AKC88560.1_Taeniopygia_guttata_Germany
putative_Rep_protein_Porcine_circovirus_2_AAC55462.1
Rep_protein_bat_associated_cyclovirus_5_ADU76999.1_bat_USA
replication-associated_protein_Pacific_flying_fox_associated_cyclovirus-3_AMH87654.1_Pteropus_tonganus_Tonga
81.5/33
replication_protein_Crucivirus_sp._WZH56856.1_Neophema_chrysogaster_Australia
replication_protein_Circovirus_sp._WRQ56666.1_Chelonia_mydas_Australia
100/99
Triatoma spp. circoviridae vOTU35
replication_protein_Circovirus_sp._WZH56888.1_Neophema_chrysogaster_Australia
replication_protein_partial_Circovirus_sp._XNO10979.1_Ixodes_ovatus_China
replication_protein_Circovirus_sp._WGL41059.1_Neophema_chrysogaster_Australia
0.7
C
Microviridae (Capsid)
Circovirus_sp._WGL41060.1_feces_Neophema_chrysogaster_Australia
Termite_gut_associated_microvirus_1_AYQ58815.1_gut_Coptotermes_formosanus_USA
Gokushovirus_WZ-2015a_ALS03762.1_Homo_sapiens_France
94/92
Microvirus_sp._WNK12593.1_Potamopyrgus_antipodarum_New_Zealand
Triatoma rubida microviridae vOTU37
Blackfly_microvirus_SF02_QCQ85111.1_whole_insect_Austrosimulium_ungulatum_New_Zealand
Apis_mellifera_associated_microvirus_8_AZL82811.1_Hemolymph_Apis_mellifera_linguistica_USA
Ariziama_microvirus_QXP07984.1_Lake_water_sample_USA
River_basin_pequenovirus_AJK28256.1_cold_methane_seep_sediment_USA
robinz_microvirus_RP_160_UDN67867.1_feces_of_robin_Petroica_australis_New_Zealand
Apis_mellifera_associated_microvirus_17_AZL82963.1_Hemolymph_Apis_mellifera_linguistica_USA
Kummerowia_striata_gokushovirus_QKI28927.1_Kummerowia_striata
96/88
Apis_mellifera_associated_microvirus_18_AZL82969.1_Apis_mellifera_linguistica_USA
0.3
D
Partitiviridae (RdRp)
Cryptosporidium_parvum_virus_1_AAC47805.1
Triatoma spp. partitiviridae vOTU30
99.9/100
Hangzhou_partitivirus_1_UHK03007.1_Rice_Field_Orthetrum_chrysostigma_China
Botryosphaeria_dothidea_partitivirus_2_UVZ34174.1_Botryosphaeria_dothidea_China
47.6/85
Hubei_partiti-like_virus_12_APG78257.1_Odonata_China
Beauveria_bassiana_partitivirus_3_QFP40245.1_Beauveria_bassiana_China
Mute_swan_feces_associated_partitiviridae_L_QUS52685.1_Cygnus_olor_fecal_material_United_Kingdom
99/99
Beihai_partiti-like_virus_1_APG78183.1_woodlouse_China
91.3/84
Triatoma indictiva partitiviridae vOTU29
Triatoma sanguisuga partitiviridae vOTU31
72/86
Hubei_partiti-like_virus_10_APG78227.1_Odonata_China
75.1/75
Triatominae partitiviridae vOTU32
84.9/88
Hubei_partiti-like_virus_11_YP_009329875.1_Odonata_China
87.5/59
Verdadero_virus_QMI58128.1_Aedes_aegypti_Mexico
Hubei_partiti-like_virus_16_APG78316.1_pillworm_China
Dragana_partiti-like_virus_QRD99907.1_Culex_modestus
Hubei_partiti-like_virus_13_APG78275.1_spiders_China
Hubei_partiti-like_virus_6_APG78276.1_spiders_China
93.6/59
Hubei_partiti-like_virus_5_APG78252.1_Odonata_China
100/100
Hubei_partiti-like_virus_4_APG78224.1_Coleoptera_China
Hubei_partiti-like_virus_1_APG78306.1_insects_China
Bactrocera_dorsalis_partiti-like_virus_1_UPT53664.1_Bactrocera_dorsalis_China
91.5/67
Hubei_partiti-like_virus_3_APG78162.1_arthropod_China
Vannostrand_partiti-like_virus_QIJ70090.1_Calliphora_vicina_pool_Australia
Atrato_Partiti-like_virus_2_QHA33902.1_Anopheles_darlingi_Colombia
Partitiviridae_sp._WAK75578.1_Other_animal_Amblyomma_varanense_China
Nanning_Parti_tick_virus_1_UYL95489.1_animal_Amblyomma_varanense_China
Hubei_partiti-like_virus_7_APG78230.1_Odonata_China
87/71
Soybean_thrips_partiti-like_virus_9_QQO81405.1_Neohydatothrips_variabilis_USA
0.5

Figure 4

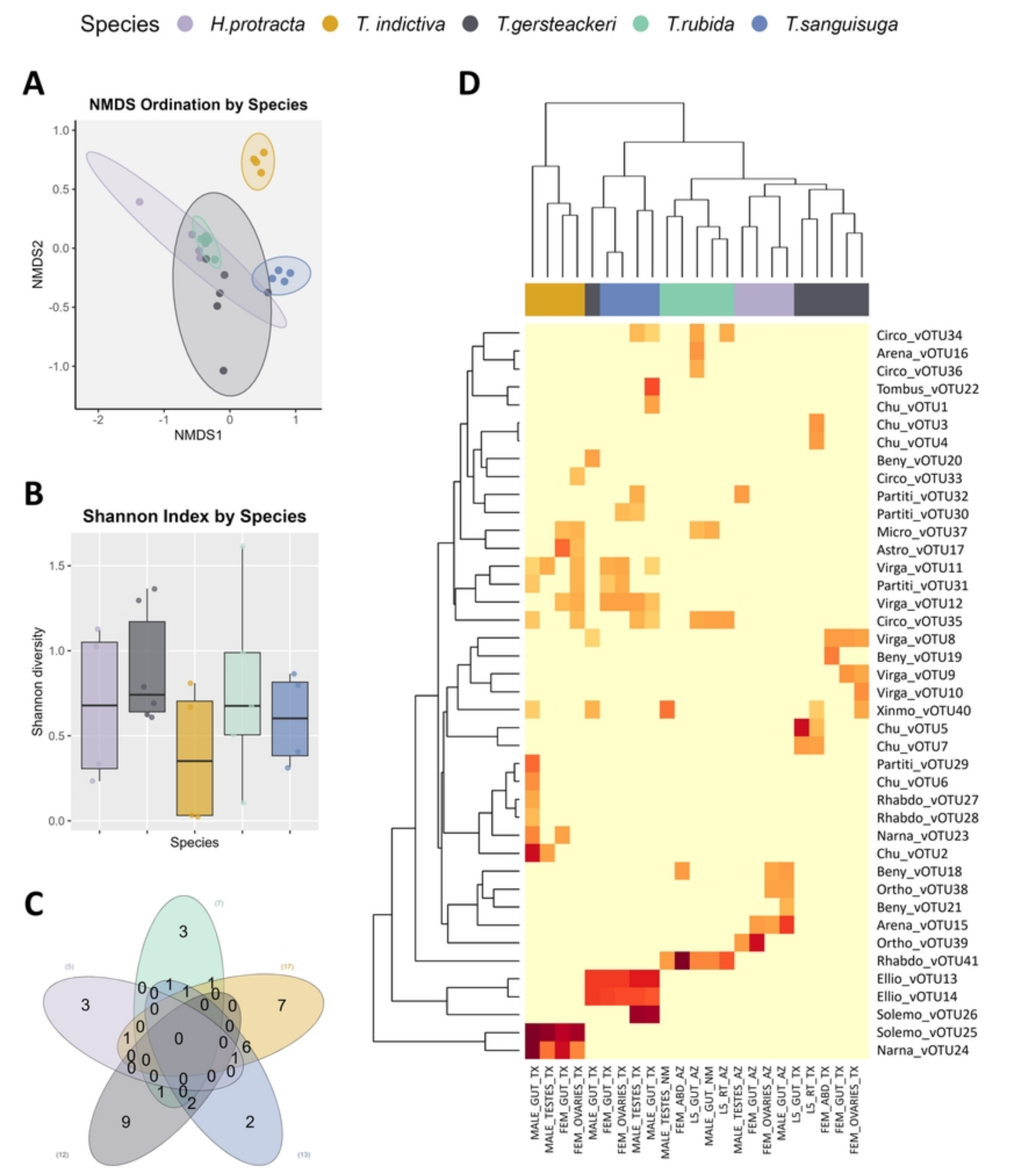
Species
H.protracta
T. indictiva
T.gersteackeri
T.rubida
T.sanguisuga
A
NMDS Ordination by Species
NMDS2
NMDS1
B
Shannon Index by Species
Shannon diversity
Species
C
D
Circo_vOTU34
Arena_vOTU16
Circo_vOTU36
Tombus_vOTU22
Chu_vOTU1
Chu_vOTU3
Chu_vOTU4
Beny_vOTU20
Circo_vOTU33
Partiti_vOTU32
Partiti_vOTU30
Micro_vOTU37
Astro_vOTU17
Virga_vOTU11
Partiti_vOTU31
Virga_vOTU12
Circo_vOTU35
Virga_vOTU8
Beny_vOTU19
Virga_vOTU9
Virga_vOTU10
Xinmo_vOTU40
Chu_vOTU5
Chu_vOTU7
Partiti_vOTU29
Chu_vOTU6
Rhabdo_vOTU27
Rhabdo_vOTU28
Narna_vOTU23
Chu_vOTU2
Beny_vOTU18
Ortho_vOTU38
Beny_vOTU21
Arena_vOTU15
Ortho_vOTU39
Rhabdo_vOTU41
Ellio_vOTU13
Ellio_vOTU14
Solemo_vOTU26
Solemo_vOTU25
Narna_vOTU24
MALE_GUT_TX
MALE_TESTES_TX
FEM_GUT_TX
FEM_OVARIES_TX
MALE_GUT_TX
FEM_GUT_TX
FEM_OVARIES_TX
MALE_TESTES_TX
MALE_GUT_TX
MALE_TESTES_NM
FEM_ABD_AZ
L5_GUT_AZ
MALE_GUT_NM
L5_RT_AZ
MALE_TESTES_AZ
FEM_GUT_AZ
FEM_OVARIES_AZ
MALE_GUT_AZ
L5_GUT_TX
L5_RT_TX
FEM_ABD_TX
FEM_GUT_TX
FEM_OVARIES_TX

Figure 5