Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses

Ci-Xiu Li, Mang Shi, Jun-Hua Tian, Xian-Dan Lin, Yan-Jun Kang, Liang-Jun Chen, Xin-Cheng Qin, Jianguo Xu, Edward C Holmes, Yong-Zhen Zhang

DOI: http://dx.doi.org/10.7554/eLife.05378

Cite as: eLife 2015;10.7554/eLife.05378

Received: 29 October 2014
Accepted: 27 January 2015
Published: 29 January 2015

Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses

Ci-Xiu Li[a1], Mang Shi[a,b1], Jun-Hua Tian[c1], Xian-Dan Lin[d1], Yan-Jun Kang[a1], Liang-Jun Chen[a], Xin-Cheng Qin[a], Jianguo Xu[a], Edward C. Holmes[a,b], Yong-Zhen Zhang[a2]

[a]State Key Laboratory for Infectious Disease Prevention and Control, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Changping, 100206, Beijing, China.
[b]Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Biological Sciences and Sydney Medical School, The University of Sydney, Sydney, NSW 2006, Australia.
[c]Wuhan Center for Disease Control and Prevention, Wuhan, 430015, Hubei Province, China
[d]Wenzhou Center for Disease Control and Prevention, Wenzhou, 325001, Zhejiang Province, China.

[1]Contributed to this work equally.

[2]Correspondence to: Dr. Yong-Zhen Zhang, State Key Laboratory for Infectious Disease Prevention and Control, National Institute of Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Changping Liuzi 5, Beijing, 102206, China. Tel: 086-10-58900782; Email: zhangyongzhen@icdc.cn.

Major subject areas: Evolution, Microbiology

Keywords: virus, evolution, phylogeny, arthropods, segmentation, negative-sense

Major Research Organisms: viruses

Abstract: Although arthropods are important viral vectors, the biodiversity of arthropod viruses, as well as the role that arthropods have played in viral origins and evolution, is unclear. Through RNA sequencing of 70 arthropod species we discovered 112 novel viruses that appear to be ancestral to much of the documented genetic diversity of negative-sense RNA viruses, a number of which are also present as endogenous genomic copies. With this greatly enriched diversity we revealed that arthropods contain viruses that fall basal to major virus groups, including the vertebrate-specific arenaviruses, filoviruses, hantaviruses, influenza viruses, lyssaviruses, and paramyxoviruses. We similarly documented a remarkable diversity of genome structures in arthropod viruses, including a putative circular form, that sheds new light on the evolution of genome organization. Hence, arthropods are a major reservoir of viral genetic diversity and have likely been central to viral evolution.

Impact statement: We document extensive genetic diversity and novel genome structures in RNA viruses from arthropods, shedding important new light on the ancestry and evolutionary history of major classes of vertebrate and plant viruses.

Introduction

Negative-sense RNA viruses are important pathogens that cause a variety of diseases in humans including influenza, hemorrhagic fever, encephalitis, and rabies. Taxonomically, those negative-sense RNA viruses described to date comprise at least eight virus families and four unassigned genera or species (King et al., 2012). Although they share (i) an homologous RNA-dependent RNA polymerase (RdRp), (ii) inverted complementary genome ends, and (iii) an encapsidated negative-sense RNA genome, these viruses display substantial diversity in terms of virion morphology and genome organization (King et al., 2012). One key aspect of genome organization is the number of distinct segments, which is also central to virus classification. Among negative-sense RNA viruses, the number of segments varies from one (order *Mononegavirales*; unsegmented) to two (family *Arenaviridae*), three (*Bunyaviridae*), three-to-four (*Ophioviridae*), and six-to-eight (*Orthomyxoviridae*), and is further complicated by differences in the number, structure, and arrangement of the encoded genes.

Despite their diversity and importance in infectious disease, the origins and evolutionary history of the negative-sense RNA viruses is largely obscure. Arthropods harbor a diverse range of RNA viruses, which are often divergent from those that infect vertebrates (Ballinger et al., 2014; Cook et al., 2013; Marklewitz et al., 2011; Marklewitz et al., 2013; Qin et al., 2014; Tokarz et al., 2014a; Tokarz et al., 2014b). However, those arthropod viruses sampled to date are generally those that have a relationship with vertebrates or are known to be agents of disease (Junglen and Drosten, 2013). To determine the extent of viral diversity harbored by arthropods, as well as their evolutionary history, we performed a systematic survey of negative-sense RNA viruses using RNA sequencing (RNA-seq) on a wide range of arthropods.

Results

Discovery of highly divergent negative-sense RNA viruses. We focused our study of virus biodiversity and evolution on 70 potential host species from four arthropod classes: Insecta, Arachnida, Chilopoda, and Malacostraca (Table 1 and Figure 1). From these samples, 16 separate cDNA libraries were constructed and sequenced, resulting in a total of 147.4 Gb of 100-base pair-end reads (Table 1). Blastx comparisons against protein sequences of negative-sense RNA virus revealed 108 distinct types of complete or nearly complete large (L) proteins (or polymerase protein 1 (PB1) in the case of orthomyxoviruses) that encode the relatively conserved RdRp (Tables 2-4). Four additional types of previously undescribed RdRp sequence (>1000 amino acids) were identified from the Transcriptome Shotgun Assembly (TSA) database. Together, these proteins exhibited an enormous diversity in terms of sequence variation and structure. Most notably, this data set of RdRp sequences is distinct from both previously described sequences and from each other, with the most divergent showing as little as 15.8% amino acid sequence identity to its closest relatives (Tables 2-4). Overall, these data provide evidence for at least 16 potentially new families and genera of negative-sense RNA viruses, defined as whose RdRp sequences shared less than 25% amino acid identity with existing taxa.

Next, we measured the abundance of these sequences as the number transcripts per million (TPM) within each library after the removal of rRNA reads. The abundance of viral transcripts calculated in this manner exhibited substantial variation (Figure 2, Tables 2-4): while the least abundant L segment (Shayang Spider Virus 3) contributed to less than 0.001% to the total non-ribosomal RNA content, the most abundant (Sanxia Water Strider Virus 1) was at a frequency of

21.2%, and up to 43.9% if we include the matching M and S segments of the virus. The remaining viral RdRp sequences fell within a range (10-1000 TPM) that matched the abundance level of highly expressed host mitochondrial genes (Figure 2).

Evolutionary history of negative-sense RNA viruses. With this highly diverse set of RdRp sequences in hand we re-examined the evolution of all available negative-sense RNA viruses by phylogenetic analysis (Figure 3; Figure 3—figure supplement 1-3). These data greatly expand the documented diversity of four viral families/orders – the *Arenaviridae*, *Bunyaviridae*, *Orthomyxoviridae*, and *Mononegavirales* – as well as of three floating genera – *Tenuivirus*, *Emaravirus*, and *Varicosavirus* (King et al., 2012). Most of the newly described arthropod viruses fell basal to the known genetic diversity in these taxa: their diversity either engulfed that of previously described viruses, as in the case of phlebovirus, nairovirus, and dimarhabovirus, or appeared as novel lineages sandwiched between existing genera or families, and hence filling in a number of phylogenetic 'gaps' (Figure 3; Figure 3—figure supplement 1-3). One important example was a large monophyletic group of newly discovered viruses that fell between the major groups of segmented and unsegmented viruses (Figure 4); we name this putative new virus family the 'Chuviridae' reflecting the geographic location in China where most of this family were identified ("Chu" is an historical term referring to large area of China encompassing the middle and lower reaches of the Yangzi River). Also of note was that some of the previously defined families no longer appear as monophyletic. For example, although classified as distinct families, the family *Arenaviridae* fell within the genetic diversity of the family *Bunyaviridae* and as a sister group to viruses of the genus *Nairovirus*. Furthermore, the floating genus *Tenuivirus* was nested within the Phlebovirus-like clade, and another floating genus, *Emaravirus*, formed a monophyletic group with the *Orthobunyavirus* and *Tospovirus* genera (Figure 3C; Figure 3—

figure supplement 2). Hence, there are important inconsistencies between the current virus classification scheme and the underlying evolutionary history of the RdRp revealed here.

A key result of this study is that the much of the genetic diversity of negative-sense RNA viruses in vertebrates and plants now appears to be contained within viruses that utilize arthropods as hosts or vectors. Indeed, it is striking that all vertebrate-specific segmented and unsegmented viruses (arenavirus, bornavirus, filovirus, hantavirus, influenza viruses, lyssavirus, and paramyxovirus) fall within the genetic diversity of arthropod-associated viruses (Figures 3 and 5). Also nested with arthropod-associated diversity were plant viruses (emaravirus, tospovirus, and tenuiviruses, nucleorhabdovirus, cytorhabdovirus, and varicosavirus) (Figures 3 and 5). Surprisingly, our phylogeny similarly placed two non-arthropod invertebrate viruses, found in nematodes (*Heterodera glycines*) and flatworms (*Procotyla fluviatilis*), within arthropod-associated diversity (Figure 3C, Figure 3—figure supplement 2), indicating that the role of non-arthropod invertebrates should be explored further. Finally, it was striking that although individual arthropod species can harbor a rich diversity of RNA viruses, many viruses seemed to be associated with different arthropod species that share the same ecological niche (Tables 2-4). Interestingly, host species in the same niche had similar viral contents that were generally incongruent with the host phylogeny (Figure 6). Such a pattern is indicative of frequent cross-species and occasional cross-genus virus transmission in the context of ecological and geographic proximity.

Diversity and evolution of virus genome organizations. The diversity of genome structures in these virus data was also striking. This can easily be documented with respect to the evolution of genome segmentation. The number of genome segments in negative-sense RNA viruses varies from one to eight. Our phylogenetic analysis revealed no particular trend for this number to

133 increase or decrease through evolutionary time (Figure 4). Hence, genome segmentation (i.e.
134 genomes with >1 segment) has clearly evolved on multiple occasions within the negative-sense
135 RNA viruses (Figure 4), such that it is a relatively flexible genetic trait. Although most
136 segmented viruses were distantly related to those with a single segment (Figure 4), close
137 phylogenetic ties were seen in other cases supporting the relatively recent evolution of multiple
138 segments, with the plant-infecting varicosavirus (two segments) and orchid fleck virus (bipartite)
139 serving as informative examples.

140 In this context it is notable that the newly discovered chuviruses fell 'between' the phylogenetic
141 diversity of segmented and the unsegmented viruses. Although monophyletic, the chuviruses
142 display a wide variety of genome organizations including unsegmented, bi-segmented, and a
143 circular form, each of which appeared multiple times in the phylogeny (Figure 4 and 7). The
144 circular genomic form, which was confirmed by 'around-the-genome' RT-PCR and by the
145 mapping of sequencing reads to the genome (Figure 7C), is a unique feature of the Chuviridae,
146 and can be distinguished from a pseudo-circular structure seen in some other negative-sense
147 RNA viruses including the family *Bunyaviridae* and the family *Orthomyxoviridae*. Furthermore,
148 this circular genomic form was also present in both segments of the segmented chuviruses
149 (Figure 7B). In addition, the chuviruses displayed a diverse number and arrangement of
150 predicted open reading frames that were markedly different from the genomic arrangement seen
151 in the order *Mononegavirales* even though these viruses are relatively closely related (Figure 4
152 and 7). In particular, the chuviruses had unique and variable orders of genes: the linear chuvirus
153 genomes began with the glycoprotein (G) gene, followed by the nucleoprotein (N) gene and then
154 the polymerase (L) gene, whereas the majority of circular chuviruses were most likely arranged
155 in the order L-(G)-N (i.e. if displayed in a linear form) as the only low coverage point throughout

the genome lay between the 5' end of N gene and the 3' end of L gene (Figure 7B). In addition, the genome organizations of the chuviruses were far more concise than those of the order *Mononegavirales*, with ORFs encoding only 2-3 major (> 20kDa) proteins (Figure 7), and hence showing more similarity to segmented viruses in this respect.

Although our phylogenetic analysis focused on the relatively conserved RdRp, in the case of segmented viruses we searched for other putative viral proteins from the assembled contigs. Accordingly, we were able to find the segments encoding matching structural proteins (mainly glycoproteins and nucleoproteins) for many of the viral RdRp sequences (Figure 8), although extensive sequence divergence prevented this in some cases. Surprisingly, M segments were apparently absent in a group of tick phleboviruses whose RdRps and nucleoproteins showed relatively high sequence similarity to Uukuniemi virus (genus *Phlebovirus*; Table 3 and Figure 8). Genomes with missing glycoprotein genes were also found in the chuviruses (Changping Tick Viruses 3 and 5, Wuhan Louse Viruses 6 and 7, Figure 7) and the unsegmented dimarhabdovirus (Taishun Tick Virus, Wuhan Tick Virus 1, Tacheng Tick Virus 6, Figure 9). Although it is possible that the glycoprotein gene may have been replaced with a highly divergent or even non-homologous sequence, we failed to find any candidate G proteins within the no-Blastx-hit set of sequences under the following criteria: (i) structural resemblance to G proteins, (ii) similar level of abundance to the corresponding RdRp and nucleoprotein genes, and (iii) comparable phylogenies or levels of divergence (among related viruses) to those of RdRps and nucleoproteins. The cause and biological significance of these seemingly "incomplete" virus genomes requires further study. Finally, it was also of interest that a virus with four segments was discovered in the horsefly pool. Although the predicted proteins of all four segments showed sequence homology to their counterparts in Tenuivirus (Falk and Tsai, 1998), this virus lacked

the ambisense coding strategy of tenuiviruses (Figure 10). While the capability of this virus to infect plants is unknown, it is possible that it represents a transitional form between plant-infecting and arthropod-specific viruses.

Novel Endogenous Virus Elements (EVEs). As well as novel exogenous RNA viruses, our metagenomic analysis also revealed a large number of potential EVEs (Katzourakis and Gifford, 2010) in more than 40 arthropod species; these resembled complete or partial genes of the major proteins – the nucleoprotein, glycoprotein and RdRp – but without fully intact genomes (Table 5). As expected given their endogenous status, most of these sequences have disrupted reading frames and many are found within transposon elements, suggesting that transposons have been central to their integration. Interestingly, in some cases, such as the putative glycoprotein gene of chuviruses, the homologous EVEs from within a family (Culicidae) or even an order (Hymenoptera) form monophyletic groups (Figure 11). However, they are unlikely to be orthologous because they do not share homologous integration sites in the host genome as determined by an analysis of flanking sequences, which in turn limited the applicability of molecular-clock based dating techniques. Furthermore, phylogenetic analyses of those EVEs shared among different host species revealed extremely complex tree topologies which do not exhibit simple matches to the host phylogeny at both the species and genera levels (Figure 11B-C). In sum, these results suggest that EVEs are relative commonplace in arthropod genomes and have been often generated by multiple and independent integration events.

Discussion

Our study suggests that arthropods are major reservoir hosts for many, if not all, of the negative-sense RNA viruses in vertebrates and plants, and hence have likely played a major role in their evolution. This is further supported by the high abundance of viral RNA in the arthropod transcriptome, as well as by the high frequencies of endogenous copies of these viruses in the arthropod genome, greatly expanding the known biodiversity of these genomic "fossils" (Cui and Holmes, 2012; Katzourakis and Gifford, 2010). The often basal position of the arthropod viruses in our phylogenetic trees is also compatible with the idea that the negative-sense RNA viruses found in vertebrates and plants ultimately have their ancestry in arthropods, although this will only be confirmed with a far wider sample of virus biodiversity.

The rich genetic and phylogenetic diversity of arthropod RNA viruses may in part reflect the enormous species number and diversity of arthropods, and that they sometimes live in large and very dense populations that provide abundant hosts to fuel virus transmission. Furthermore, arthropods are involved in almost all ecological guilds and actively interact with other eukaryotes, including animals, plants and fungi, such that it is possible that they serve as both sources and sinks for viruses present in the environment. In addition, not only were diverse viruses present, but they were often highly abundant. For example, in the pool containing twelve individuals (representing two species) from the Gerridae (Water striders) collected at the same site, we identified at least five negative-sense RNA viruses whose TPM values are well above 100, and where the viral RNA collectively made up more than 50% of the host total RNA (rRNA excluded). Determining why arthropods are able to carry such a large viral diversity and at such frequencies clearly merits further investigation.

The viruses discovered here also exhibited a huge variation in level of abundance. It is possible that this variation is in part due to the stage or severity of infection in individual viruses, and may

be significantly influenced by the process of pooling, since most of our libraries contain an uneven mixture of different host species or even genera. In addition, it is possible that some low abundance viruses may in fact be derived from other eukaryotic organisms present in the host sampled, such as undigested food or prey, gut micro flora, and parasites. Nevertheless, since the majority of the low abundance viruses appear in the same groups as the highly abundant ones in our phylogenetic analyses, these viruses are most likely associated with arthropods.

Viral infections in vertebrates and plants can be divided into two main categories: (i) arthropod-dependent infections, in which there is spill-over to non-arthropods but where continued virus transmission still requires arthropods, and (ii) arthropod-independent infections, in which the virus has shifted its host range to circulate among vertebrates exclusively (Figure 12). The first category of infections is often associated with major vector-borne diseases (Zhang et al., 2012; Zhang et al., 2011). Given the biodiversity of arthropod viruses documented here, it seems likely that arthropod-independent viruses were ultimately derived from arthropod-dependent infections, with subsequent adaptation to vertebrate-only transmission (Figure 12).

One of the most notable discoveries was that of a novel family, the Chuviridae. The identification of this diverse virus family provides a new perspective on the evolutionary origins of segmented and unsegmented viruses. In particular, the chuviruses occupy a phylogenetic position that is in some sense ‘intermediate’ between the segmented and unsegmented negative-sense RNA viruses, and display genomic features of both. Indeed, our phylogenetic analysis reveals that genome segmentation has evolved multiple times within the diversity of chuviruses (Figure 7), such that this trait appears to be more flexible than previously anticipated. In addition, the majority of the chuviruses possess circular genomes. To date, the only known circular RNA virus is (hepatitis) deltavirus, although this potentially originated from the human genome

(Salehi-Ashtiani et al., 2006) and requires hepatitis B virus for successful replication. As such, the chuviruses may represent the first report of autonomously replicating circular RNA viruses, which opens up an important line of future research.

Our results also provide insights into the evolution of genome segmentation. Within the Bunya-arena-like viruses (Figure 3C and 4), the three-segment structure is the most common, with the viral polymerase, nucleoprotein, and surface glycoproteins present on different segments. Notably, our phylogenetic analysis seemingly revealed independent occurrences of both increasing (Tenuivirus and Emaravirus) and decreasing (Arenavirus) of segment numbers from the three-segment form (Figure 4). Independent changes of genome segmentation numbers are also observed in the mononegavirales-like viruses (Figure 4) and, more frequently, in the chuviruses (Figure 7A). Consequently, the number of genome segments appears to be a relatively flexible trait at a broad evolutionary scale, although the functional relevance of these changes remains unclear. While the segmented viruses (bunya-arenaviruses, orthomyxoviruses, and ophioviruses) appear to be distinct from the largely unsegmented mononegavirales-like viruses in our phylogenetic analysis, this may be an artifact of under-sampling, especially given that only a tiny fraction of eukaryotes have been sampled to date. With a wider sample of eukaryotic viruses it will be possible to more accurately map changes in segment number onto phylogenetic trees and in so doing come to a more complete understanding of the patterns and determinants of the evolution of genome segmentation.

In sum, our results highlight the remarkably diversity of arthropods viruses. Because arthropods interact with a wide range of organisms including vertebrate animal and plants, they can be seen as the direct or indirect source of many clinically or economically important viruses. The viral genetic and phenotypic diversity documented in arthropods here therefore provides a new

perspective on fundamental questions of virus origins, diversity, host range, genome evolution, and disease emergence.

Materials and Methods

Sample collection. Between 2011 and 2013 we collected 70 species of arthropods from various locations in China (Table 1). Among these, ticks were either directly picked from wild and domestic animals, or captured using a tick drag-flag method; mosquitoes were trapped by light-traps; common flies were captured by fly paper; horseflies were picked from infested cattle; bed bugs and cockroaches were trapped indoors; louse flies were plucked from the skin of bats; millipedes were picked up from the ground; spiders were collected from their webs; water striders were captured using hand nets from river surfaces, and crabs and shrimps were bought (alive) from local fisherman. In addition, three pools of mixed insect samples (Table 1) were collected from a rural area adjacent to rice fields (Insect Mix 1), from a lakeside (Insect Mix 3), and from a mountainous area near Wuhan (Insect Mix 4). After brief species identification by experienced field biologists, these samples were immediately stored in liquid nitrogen and were later put on dry ice for shipment to our laboratory.

Total RNA extraction. The specimens were first grouped into several units (Table 1). Depending on the size of specimens, one unit could include from 1 to 20 individual arthropods belonging to the same species and sampling location. These units were first washed with phosphate-buffered saline (PBS) three times before homogenized with the Mixer mill MM400 (Restsch). The resultant homogenates were then subjected to RNA extraction using TRIzol LS

reagent (Invitrogen). After obtaining the aqueous phase containing total RNA, we performed purification steps from the E.Z.N.A Total RNA Kit (OMEGA) according to the manufacturer's instructions. The concentration and quality of final extractions were examined using a ND-1000 UV Spectrophotometer (Nanodrop). Based on host types and/or geographic locations, these extractions were further merged into 16 pools for RNA-seq library construction and sequencing (Table 1).

Species identification. To verify the field species identification, we took a proportion of the homogenates from each specimen or specimen pool for genomic DNA extraction using E.Z.N.A. DNA/RNA Isolation Kit (OMEGA). Two genes were used for host identification: the partial 18S rRNA gene (~ 1100nt) which was amplified using primer pairs 18S#1 (5'-CTGGTGCCAGCGAGCCGCGGYAA-3') and 18S#2RC (5'-TCCGTCAATTYCTTTAAGTT-3'), and partial COI gene (~ 680nt) using primer pairs LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA -3'). PCR reactions were performed as described previously (Folmer et al., 1994; Machida and Knowlton, 2012). For taxonomic determination, the resulting sequences were compared against the nt database as well as with all COI barcode records on the Barcode of Life Data Systems (BOLD).

RNA-seq sequencing and reads assembly. Total RNA was subjected to a slightly modified RNA-seq library preparation protocol to that provided by Illumina. Briefly, following DNase I digestion, total RNA was subjected to an rRNA removal step using Ribo-Zero™ Magnetic Gold Kit (Epidemiology). The remaining RNA was then fragmented, reverse-transcribed, ends

315 repaired, dA-tailed, adaptor ligated, purified, and quantified with Agilent 2100 Bioanalyzer and
316 ABI StepOnePlus Real-Time PCR System. Pair-end (90bp or 100bp) sequencing of the RNA
317 library was performed on the HiSeq 2000 platform (Illumina). All library preparation and
318 sequencing steps were performed by BGI Tech (Shenzhen, China). The resulting sequencing
319 reads were quality trimmed and assembled *de novo* using the Trinity program (Grabherr et al.,
320 2011). All sequence reads generated in this study were uploaded onto NCBI Sequence Read
321 Achieve (SRA) database under the BioProject accession SRP051790.

322

323 **Discovery of target virus sequences**. The assembled contigs were translated and compared
324 (using Blastx) to reference protein sequences of all negative-sense RNA viruses. Sequences
325 yielding e-values larger than $1E^{-5}$ were retained and compared to the entire nr database to
326 exclude non-viral sequences. The resulting viral sequences were merged by identifying
327 unassembled overlaps between neighboring contigs or within a scaffold using the SeqMan
328 program implemented in the Lasergene software package v7.1 (DNAstar, Madison, WI). To
329 prevent missing highly divergent viruses, the newly found viral sequences were included in the
330 reference protein sequences for a second round of Blastx.

331

332 **Sequence confirmation and repairing by Sanger methods**. For each potential viral sequence,
333 we first used nested RT-PCR to examine which unit contained the target sequence, utilizing
334 primers designed based on the deep-sequencing results. In the case of segmented viruses this
335 information was also used to determine whether and which of the segments recovered from the
336 pool belonged to the same virus. We next designed overlapping primers to verify the sequence
337 obtained from the deep sequencing and assembly processes. Based on the verified sequences, we

338 determined the sequencing depth and coverage by mapping reads to target sequences using
339 bowtie2 (Langmead and Salzberg, 2012). All virus genome sequences generated in this study
340 have been deposited in the GenBank database under accession numbers KM817593-KM817764.

341

342 **Quantification of relative transcript abundances.** Before quantification, we first removed the
343 rRNA reads from the data sets to prevent any bias due to the unequal efficiency of rRNA
344 removal steps during library preparation. To achieve this, we blasted the Trinity assembly results
345 against the SILVER rRNA database (Quast et al., 2013), and then used the resulting rRNA
346 contigs as a template for mapping using BOWTIE2 (Langmead and Salzberg, 2012). The
347 remaining reads from each library were then mapped on to the assembled transcripts and
348 analyzed with RSEM (Li et al., 2010), using the run_RSEM_align_n_estimate.pl scripts
349 implemented in the Trinity program (Grabherr et al., 2011). The relative abundance of each
350 transcript is presented as transcripts per million (TPM) which corrects for the total number of
351 reads as well as for transcript length (Li et al., 2010).

352

353 **Genome walking**. Some of the sequences obtained were substantially shorter than expected. To
354 obtain longer sequences, we used a Genome walking kit (TaKaRa). Briefly, three gene-specific
355 primers close to the end of the known sequence were designed. RNA from positive samples was
356 used as input for reverse transcription primed by random primer N6. TAIL-PCR (thermal
357 asymmetric interlaced PCR) was performed according to the manufacturer's protocol. The
358 cDNA was used as a template for PCR with specific primers and the manufacturer-supplied
359 degenerate primers. After three rounds of amplification, the products were analyzed on 1.0%
360 agarose gels, and single fragments were recovered from the gels and purified using an agarose

gel DNA extraction kit (TaKaRa). The purified products were then ligated into pMD19-T vector (TaKaRa) which contains the gene for ampicillin resistance. The vector was transformed into DH5α cells, which were spread on agar plates and incubated overnight at 37°C. A total of 10 clones were randomly selected and sequenced using M13 primers on ABI 3730 genetic analyzer (Applied Biosystems).

Determination of genome/segment termini. The extreme 5' sequences were recovered by performing a 5'-Full RACE kit with TAP (TaKaRa) according to the manufacturer's protocol. Briefly, two gene-specific primers close to the end of the known sequence were designed. The 5' end of RNA was ligated to the 5'RACE adaptor (without 5' end dephosphorylating and decapping) and then reverse-transcribed using random 9 mers. The resulting cDNA was used as a template for nested PCR with 5′ RACE primers provided by the kit and gene-specific reverse primers. The PCR products were separated on an agarose gel, cloned into pMD19-T cloning vector, and subsequently sequenced.

The extreme 3' sequences were recovered by performing a 3'-full RACE Core Set with PrimeScript RTase (TaKaRa) according to the manufacturer's protocols. Because the RNA template lacks a polyadenylated tail, a Poly(A) Tailing Kit (Applied Biosystems) was used to add this to the RNAs prior to first-strand 3'-cDNA synthesis. 20μL of the Poly(A)-tailing reaction mixture was prepared according to the manufacturer's instructions and was incubated at 37°C for 1 hr before reverse transcription using PrimeScript Reverse Transcriptase. The cDNA was then amplified by nested PCR using the 3′ RACE primers provided by the kit and gene-specific reverse primers. The PCR products were separated on agarose gels, cloned into pMD19-T cloning vector, and subsequently sequenced. The 5' and 3' ends of the genome fragment were

384 also determined by RNA circularization. RT-PCR amplification was performed across the ligated
385 termini and the resulting PCR products were subsequently cloned and sequenced.

386

387 **Phylogenetic analyses**. Potential viral proteins identified from this study were aligned with their
388 corresponding homologs of reference negative-sense RNA viruses using MAFFT version 7 and
389 employing the E-INS-i algorithm (Katoh and Standley, 2013). The sequence alignment was
390 limited to conserved domains, with ambiguously aligned regions removed using TrimAl
391 (Capella-Gutierrez et al., 2009). The final alignment lengths were 224 amino acids (aa), 412aa,
392 727aa, and 364aa for data sets of overall, bunya-arena-like, mononega-like, and orthomyxo-like
393 data sets, respectively. Phylogenetic trees were inferred using the maximum likelihood method
394 (ML) implemented in PhyML version 3.0 (Guindon and Gascuel, 2003), with the WAG+Γ
395 amino acid substitution model and a Subtree Pruning and Regrafting (SPR) topology searching
396 algorithm. Phylogenetic trees were also inferred using a Bayesian method implemented in
397 MrBayes version 3.2.2 (Ronquist and Huelsenbeck, 2003), with the same substitution model as
398 used in ML tree inference. In the MrBayes analyses, we used two simultaneous runs of Markov
399 chain Monte Carlo sampling, and the runs were terminated upon convergence (standard
400 deviation of the split frequencies <0.01). The phylogeny was subsequently summarized from
401 both runs with an initial 10% of trees discarded as burn-in.

402

403 **Prediction of protein domains and functions**. For each of the putative viral protein sequences,
404 we used TMHMM v2.0 (http://www.cbs.dtu.dk/services/TMHMM/) to predict the
405 transmembrane domains, SignalP v4.0 (http://www.cbs.dtu.dk/serv-ices/SignalP/) to determine

signal sequences, and NetNGlyc v1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) to identify N-linked glycosylation sites. For some of the highly divergent viruses belonging to the Mononegavirales and the Chuviridae, a protein was regarded as a potential glycoprotein if it contained (i) a N-terminal signal domain, (ii) a C-terminal transmembrane domain, and (iii) glycosylation sites in cytoplasmic domains.

Identification and characterization of endogenous viruses. Endogenous copies of the exogenous negative-sense RNA viruses newly described here were detected using the tBlastn algorithm against arthropod genomes available in the Reference Genomic Sequences Database (refseq_genomic) and Whole Genome Shotgun Database (WGS) in GenBank, and using viral amino acid sequences as queries. The threshold for match was set to 1e-05 for the e-value and 50 amino acids for matched length. The query process was reversed for each potential endogenous virus to determine their corresponding phylogenetic group. Orthologous insertion events were determined by examining flanking gene sequences. Sequence alignment and phylogenetic analyses were carried out as described above.

Characterization of bi-segmented viruses in the Chuviridae. Within the Chuviridae, Wuhan Louse Fly Virus 6 and 7, Wenzhou Crab Virus 2, Lishi Spider Virus 1, and Wuchang Cockroach Virus 3 possessed bi-segmented genomes. Both segments were discovered using Blastx against pools of predicted proteins from unsegmented chuvirus or mononegavirales sequences. To determine that these sequences were indeed from separate segments, we performed all combinations of head-to-tail RT-PCR which allowed us to ascertain whether the sequence fragments came from a single genome. Furthermore, checking sequencing depth can help

eliminate the possibility of separate contigs being generated due to inadequate sequencing coverage. To prove that a pair of segments belonged to the same virus, we checked; (i) sequencing depth for both segments, (ii) the presence of conserved regulatory sequences at non-coding regions of the genome, (iii) whether there is match for PCR-positive units, and (iv) the phylogenetic positions of the different viral proteins (Figure 7A).

Characterization of a circular genome form within the Chuviridae. The circular genome organization within the Chuviridae was identified after we found that their genome sequences were "over assembled" (i.e. generating contigs that contained more than one genome connected head-to-tail). This circular genomic form was also observed in both segments of the segmented chuviruses (Figure 7B). In addition, RT-PCR and sequencing over the entire genome did not reveal any break-points. As a control, the same protocol failed to connect the genome termini within the Mononegavirales, suggesting the circular genomic form is unique to the chuviruses. To further validate that these genomes are circular, we mapped the high-throughput sequencing reads to these assembled genomes. The coverage and depth was adequate throughout the genome with the exception of one location upstream to the 3' end of the ORF encoding RdRp (Figure 7C). This genomic location had only 0-20 X coverage depending on the virus, although all RT-PCRs were successful across this location. Interestingly, sequencing of the cloned PCR products revealed extensive sequence variation (i.e. insertions and deletions) (Figure 7C), which is the likely cause of the low sequence coverage in this location. Collectively, these data provide strong evidence for circular genomes in the chuviruses, although this does not exclude the potential presence of linear genomic forms.

References

Ballinger MJ, Bruenn JA, Hay J, Czechowski D, Taylor DJ. 2014. Discovery and evolution of bunyavirids in arctic phantom midges and ancient bunyavirid-like sequences in insect genomes. *Journal of virology* **88**:8783-8794. doi: 10.1128/JVI.00531-14.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**:1972-1973. doi: 10.1093/bioinformatics/btp348.

Cook S, Chung BY, Bass D, Moureau G, Tang S, McAlister E, Culverwell CL, Glucksman E, Wang H, Brown TD, Gould EA, Harbach RE, de Lamballerie X, Firth AE. 2013. Novel virus discovery and genome reconstruction from field RNA samples reveals highly divergent viruses in dipteran hosts. *PloS one* **8**:e80720. doi: 10.1371/journal.pone.0080720.

Cui J, Holmes EC. 2012. Endogenous RNA viruses of plants in insect genomes. *Virology* **427**:77-79. doi: 10.1016/j.virol.2012.02.014.

Falk BW, Tsai JH. 1998. Biology and molecular biology of viruses in the genus Tenuivirus. *Annual review of phytopathology* **36**:139-163. doi: 10.1146/annurev.phyto.36.1.139.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular marine biology and biotechnology* **3**:294-299. doi:

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* **29**:644-652. doi: 10.1038/nbt.1883.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic biology* **52**:696-704. doi:

Junglen S, Drosten C. 2013. Virus discovery and recent insights into virus diversity in arthropods. *Current opinion in microbiology* **16**:507-513. doi: 10.1016/j.mib.2013.06.005.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* **30**:772-780. doi: 10.1093/molbev/mst010.

Katzourakis A, Gifford RJ. 2010. Endogenous viral elements in animal genomes. *PLoS genetics* **6**:e1001191. doi: 10.1371/journal.pgen.1001191.

King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ. 2012. Virus Taxonomy: 9th Report of the International Committee on Taxonomy of Viruses: Elsevier Academic Press.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**:357-359. doi: 10.1038/nmeth.1923.

Li B, Ruotti V, Stewart RM, Thomson JA, Dewey CN. 2010. RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics* **26**:493-500. doi: DOI 10.1093/bioinformatics/btp692.

Machida RJ, Knowlton N. 2012. PCR primers for metazoan nuclear 18S and 28S ribosomal DNA sequences. *PloS one* **7**:e46180. doi: 10.1371/journal.pone.0046180.

Marklewitz M, Handrick S, Grasse W, Kurth A, Lukashev A, Drosten C, Ellerbrok H, Leendertz FH, Pauli G, Junglen S. 2011. Gouleako virus isolated from West African mosquitoes constitutes a proposed novel genus in the family Bunyaviridae. *Journal of virology* **85**:9227-9234. doi: 10.1128/JVI.00230-11.

Marklewitz M, Zirkel F, Rwego IB, Heidemann H, Trippner P, Kurth A, Kallies R, Briese T, Lipkin WI, Drosten C, Gillespie TR, Junglen S. 2013. Discovery of a unique novel clade of mosquito-associated bunyaviruses. *Journal of virology* **87**:12850-12865. doi: 10.1128/JVI.01862-13.

Qin XC, Shi M, Tian JH, Lin XD, Gao DY, He JR, Wang JB, Li CX, Kang YJ, Yu B, Zhou DJ, Xu J, Plyusnin A, Holmes EC, Zhang YZ. 2014. A tick-borne segmented RNA virus contains genome segments derived from unsegmented viral ancestors. *Proceedings of the National Academy of Sciences of the United States of America* **111**:6744-6749. doi: 10.1073/pnas.1324194111.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* **41**:D590-596. doi: 10.1093/nar/gks1219.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**:1572-1574. doi:

Salehi-Ashtiani K, Luptak A, Litovchick A, Szostak JW. 2006. A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science* **313**:1788-1792. doi: 10.1126/science.1129308.

Tokarz R, Sameroff S, Leon MS, Jain K, Lipkin WI. 2014a. Genome characterization of Long Island tick rhabdovirus, a new virus identified in Amblyomma americanum ticks. *Virology journal* **11**:26. doi: 10.1186/1743-422X-11-26.

Tokarz R, Williams SH, Sameroff S, Sanchez Leon M, Jain K, Lipkin WI. 2014b. Virome Analysis of Amblyomma americanum, Dermacentor variabilis, and Ixodes scapularis Ticks Reveals Novel Highly Divergent Vertebrate and Invertebrate Viruses. *Journal of virology* **88**:11480-11492. doi: 10.1128/JVI.01858-14.

Zhang YZ, Zhou DJ, Qin XC, Tian JH, Xiong Y, Wang JB, Chen XP, Gao DY, He YW, Jin D, Sun Q, Guo WP, Wang W, Yu B, Li J, Dai YA, Li W, Peng JS, Zhang GB, Zhang S, Chen XM,

524 Wang Y, Li MH, Lu X, Ye C, de Jong MD, Xu J. 2012. The ecology, genetic diversity, and
525 phylogeny of Huaiyangshan virus in China. *Journal of virology* **86**:2864-2868. doi:
526 10.1128/JVI.06192-11.

527 Zhang YZ, Zhou DJ, Xiong Y, Chen XP, He YW, Sun Q, Yu B, Li J, Dai YA, Tian JH, Qin XC,
528 Jin D, Cui Z, Luo XL, Li W, Lu S, Wang W, Peng JS, Guo WP, Li MH, Li ZJ, Zhang S, Chen C,
529 Wang Y, de Jong MD, Xu J. 2011. Hemorrhagic fever caused by a novel tick-borne Bunyavirus
530 in Huaiyangshan, China. *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi*
531 **32**:209-220. doi:
532

533

534 **Acknowledgments**

535 This study was supported by National Natural Science Foundation of China (Grants 81290343,
536 81273014), the 12th Five-Year Major National Science and Technology Projects of China
537 (2014ZX10004001-005). ECH is funded by an NHMRC Australia Fellowship (AF30). The
538 authors sincerely thank Xiu-Nian Diao (Veterinary Station, Jiulingtuan of Wushi, Bole, Xinjiang
539 Uygur Autonomous Region, China) and Ming-Hui Chen (Veterinary Station, Emin, Jiushi,
540 Xinjiang Uygur Autonomous Region, China) for their assistance in sampling.

541

542

Figure legends

Figure 1. Host component of each pool used in the RNA-seq library construction and sequencing. The taxonomic units in the tree correspond to the unit samples used in the RNA extraction. Species or genus information is marked to the left of the tree.

Figure 2. Abundance level (transcripts per million – TPM) of the RdRp genes from the negative-sense RNA viruses detected in this study. Abundance is calculated after the removal of ribosomal RNA reads. As a comparison, we show the abundance of the two well characterized (positive-sense) RNA viruses: Japanese encephalitis virus and Gill-associated virus found in the Mosquito-Hubei and Shrimp libraries, respectively, as well as the range of abundance of host mitochondrial COI genes in these same multi-host libraries.

Figure 3. Evolutionary history of negative-sense RNA viruses based on RdRp. This is initially displayed in an unrooted maximum likelihood (ML) tree including all major groups of negative-sense RNA viruses (A). Separate and more detailed ML phylogenies are then shown for the Orthomyxoviridae-like (B), Bunya-Arenaviridae-like (C), and Mononegavirales-like viruses (D). In all the phylogenies, the RdRp sequences described here from arthropods are either shaded purple or marked with solid grey circles. The names of previously defined genera/families are labeled to the right of the phylogenies. Based on their host types, the branches are shaded red (vertebrate-specific), yellow (vertebrate and arthropod), green (plant and arthropod), blue (non-arthropod invertebrates) or black (arthropod only). For clarity, statistical supports (i.e. approximate likelihood-ratio test (aLRT) with Shimodaira-Hasegawa-like procedure / posterior probabilities) are shown for key internal nodes only.

Figure 3—figure supplement 1-1. A fully labelled ML phylogeny for Orthomyxoviridae-like viruses. The phylogeny is reconstructed using RdRp alignments. Statistical support from the approximate likelihood-ratio test (aLRT) is shown on each node of the tree. The names of the viruses discovered in this study are shown in red. The names of reference sequences, which contain both the GenBank accession number and the virus species name, are shown in black. The names of previously defined genera/families are shown to the right of the phylogenies.

Figure 3—figure supplement 1-2. A fully labelled ML phylogeny for Bunya-Arenaviridae-like viruses. The phylogeny is reconstructed using RdRp alignments. Statistical support from the aLRT is shown on each node of the tree. The names of the viruses discovered in this study are shown in red. The names of reference sequences, which contain both the GenBank accession number and the virus species name, are shown in black. The names of previously defined genera/families are shown to the right of the phylogenies.

Figure 3—figure supplement 1-3. A fully labelled ML phylogeny for Mononegavirales-like viruses. The phylogeny is reconstructed using RdRp alignments. Statistical support from the aLRT is shown on each node of the tree. The names of the viruses discovered in this study are shown in red. The names of reference sequences, which contain both the GenBank accession number and the virus species name, are shown in black. The names of previously defined genera/families are shown to the right of the phylogenies.

588 **Figure 4**. The unrooted ML phylogeny based on RdRp showing the topological position of
589 segmented viruses within the genetic diversity of negative-sense RNA viruses. The segmented
590 viruses are labeled with segment numbers and shaded red. The unsegmented viruses are shaded
591 green. The Chuviridae, which exhibit a wide variety of genome organizations, are shaded cyan.
592 Three major types of putative chuvirus genomes (circular, circular and segmented, and linear)
593 are shown in the right panel and are annotated with predicted ORFs: putative RdRp genes are
594 shaded blue, putative glycoprotein genes are shaded orange, and the remaining ORFs are shaded
595 grey.

596

597 **Figure 5**. The unrooted ML phylogeny of negative-sense RNA viruses (RdRp) with the common
598 names of the principle arthropod hosts analyzed in this study indicated. Vertebrate-specific
599 viruses are shaded red, those infecting both vertebrates and arthropods (or with unknown
600 vectors) are shaded yellow, those infecting both plants and arthropods are shaded green, those
601 infecting non-arthropod invertebrates are shaded blue, and the remainder (arthropod only) are
602 shaded black.

603

604 **Figure 6**. Phylogenetic congruence between viruses (M segments) and hosts, including (A)
605 Wuhan Horsefly Virus, (B) Wuhan Fly Virus 1, (C) Wuhan Mosquito Virus 2, and (D) Wuhan
606 Mosquito Virus 1. Different host species/genera are distinguished with different colors, which
607 are then mapped onto virus phylogeny to assess the phylogenetic congruence. ML phylogenetic
608 trees were inferred in all cases.

609

610 **Figure 7**. The differing genome organizations in the Chuviridae. (A) ML trees of three main
611 putative proteins conserved among the chuviruses. Viruses with circular genomes (Type I) are
612 shaded blue, while those with segmented genomes (Type II) are shaded red. (B) Structures of all
613 complete chuvirus genomes. Circular genomes are indicated with the arrow (blue) situated at the
614 3' end, and the genome is drawn in a linear form for ease of comparison only, being broken at
615 the region of variable sequence (refer to the materials and methods). (C) An example showing
616 mapping of sequencing reads to the circular chuvirus genome. The template for mapping
617 contains two genomes connected head-to-tail. The two boxes magnify the genomic region
618 containing abundant sequence variation.
619
620 **Figure 8**. Genome structures of segmented negative-sense RNA viruses. Predicted viral proteins
621 homologous to known viral proteins are shown and colored according to their putative functions.
622 The numbers below each ORF box give the predicted molecular mass.
623
624 **Figure 9**. Genome structures of unsegmented negative-sense RNA viruses. Predicted ORFs
625 encoding viral proteins with > 10kDa molecular mass are shown and colored according to their
626 putative functions. The numbers below each ORF box give the predicted molecular mass.
627
628 **Figure 10**. Comparison of the genome structure of a potential tenui-like virus from horsefly with
629 a prototype tenuivirus (Rice grassy stunt virus) genome.
630
631 **Figure 11.** ML phylogeny of EVEs based on the glycoprotein of chuviruses in the context of
632 exogenous members of this family (A), with subtrees magnified for (B) the Culicidae clade and

633 (C) the Hymenoptera clade. The EVEs used in the phylogeny covered the complete or near
634 complete length of the glycoprotein gene, and are shown in red and labeled according to host
635 taxonomy in the overall tree. For clarity, monophyletic groups are collapsed based on the host
636 taxonomy. Only bootstrap values >70% are shown.
637
638 **Figure 12.** Transmission of negative-sense RNA viruses in arthropods and non-arthropods.
639 Three types of transmission cycle are shown: (i) those between arthropods and plants are shaded
640 green; (ii) those between arthropods and vertebrates are shaded yellow; and (iii) those that are
641 vertebrate-only are shaded red. Viruses associated with each transmission type are also indicated.
642

643 **Tables**

644 **Table 1**. Host and geographic information and data output for each pool of arthropod samples
645

Pool	No of unit	Order	Species	Locations	Data generated (bases)
Mosquitos - Hubei	24	Diptera	*Aedes sp*, *Armigeres subalbatus*, *Anopheles sinensis*, *Culex quinquefasciatus*, *Culex tritaeniorhynchus*	Hubei	26,606,799,000
Mosquitos - Zhejiang	26	Diptera	*Aedes albopictus*, *Armigeres subalbatus*, *Anopheles paraliae*, *Anopheles sinensis*, *Culex pipiens*, *Culex sp*, *Culex tritaeniorhynchus*,	Zhejiang	7,233,954,480
True flies	24	Diptera	*Atherigona orientalis*, *Chrysomya megacephala*, *Lucilia sericata*, *Musca domestica*, *Sarcophaga dux*, *S. peregrina*, *S. sp*	Hubei	6,574,954,320
Horseflies	24	Diptera	unidentified *Tabanidae* (5 species)	Hubei	8,721,642,060
Cockroaches	24	Blattodea	*Blattella germanica*	Hubei	6,182,028,000
Water striders	12	Hemiptera	unidentified *Gerridae* (2 species)	Hubei	3,154,714,200
Insects mix 1	6	Diptera, Coleoptera, Lepidoptera, Neuroptera	*Abraxas tenuisuffusa*, *Hermetia illucens*, unidentified *Chrysopidae*, unidentified *Coleoptera*, *Psychoda alternata*, unidentified *Diptera*, unidentified *Stratiomyidae*	Zhejiang	7,745,172,660
Insects mix 2	4	Diptera, Hemiptera	unidentified *Hippoboscidae* (2 species), *Cimex hemipterus*	Hubei	5,916,431,520
Insects mix 3 (insect near water)	10	Odonata, Hemiptera, Hymenoptera, Isopoda	*Pseudothemis zonata*, unidentified *Nepidae* (2 species), *Camponotus japonicus*, *Diplonychus sp*, *Asellus sp*	Hubei	11,973,368,200
Insects mix 4 (insect in the mountain)	12	Diptera, Orthoptera, Odonata, Hymenoptera, Hemiptera	*Psychoda alternata*, *Velarifictorus micado*, *Crocothemis servilia*, unidentified *Phoridae*, unidentified *Lampyridae*, *Aphelinus sp*, *Hyalopterus pruni*, *Aulacorthum magnolia*,	Hubei	6,882,491,800
Ticks	16	Ixodida	*Dermacentor marginatus*, *Dermacentor sp*, *Haemaphysalis doenitzi*, *H. longicornis*, *H. sp*, *H. formosensis*, *Hyalomma asiaticum*, *Rhipicephalus microplus*, *Argas miniatus*	Hubei, Zhejiang, Beijing, Xinjiang	24,708,479,580
Ticks Hyalomma asiaticum	1	Ixodida	*Hyalomma asiaticum*	Xinjiang	2,006,000,100
Spiders	32	Araneae	*Neoscona nautica*, *Parasteatoda tepidariorum*, *Plexippus setipes*, *Pirata sp*, unidentified *Araneae*	Hubei	11,361,912,300
Shrimps	48	Decapoda	*Exopalaemon carinicauda*, *Metapenaeus sp*, *Solenocera crassicornis*, *Penaeus monodon*, *Litopenaeus vannamei*	Zhejiang	5,365,359,900
Crabs and barnacles	35	Decapoda, Scalpelliformes	*Capitulum mitella*, *Charybdis hellerii*, *C. japonica*, *Uca arcuata*	Zhejiang	5,833,269,360
Millipedes	12	Polydesmida	unidentified *Polydesmidae* (2 species)	Hubei, Beijing	7,176,702,400

646

647 **Table 2**. Mononegavirales-related RdRp sequences discovered in this study

Virus name	Length of RdRp	Classification	Pool	Abundance	Putative arthropod host	Closest relative (aa identity)
Bole Tick Virus 3	2155	chuvirus	ticks	202.35	*Hyalomma asiaticum*	Midway virus (17.1%)
Changping Tick Virus 2	2156	chuvirus	ticks	185.73	*Dermacentor sp*	Midway virus (17.6%)
Changping Tick Virus 3	2209	chuvirus	ticks	41.80	*Dermacentor sp*	Midway virus (16.5%)
Lishi Spider Virus 1	2180	chuvirus	spiders	5.82	*Parasteatoda tepidariorum*	Midway virus (16.9%)
Shayang Fly Virus 1	2459	chuvirus	true flies	8.99	*Atherigona orientalis*	Maize mosaic virus (16.8%)
Shuangao Fly Virus 1	2097	chuvirus	insect mix 1	23.63	unidentified *Diptera*	Lettuce big-vein associated virus (16.3%)
Shuangao Insect Virus 5	2291	chuvirus	insect mix 1	209.31	unidentified *Diptera*, *Abraxas tenuisuffusa*, unidentified *Chrysopidae*	Potato yellow dwarf virus (16.3%)
Shuangao Lacewing Virus	2145	chuvirus	insect mix 1	44.48	unidentified *Chrysopidae*	Potato yellow dwarf virus (16.8%)
Tacheng Tick Virus 4	2101	chuvirus	ticks	137.22	*Argas miniatus*	Midway virus (17.5%)
Tacheng Tick Virus 5	2201	chuvirus	ticks	276.32	*Dermacentor marginatus*	Midway virus (16.8%)
Wenzhou Crab Virus 2	2208	chuvirus	crabs and barnacles	4054.25	*Charybdis japonica*, *Charybdis lucifera*, *Charybdis hellerii*	Midway virus (15.8%)
Wenzhou Crab Virus 3	2077	chuvirus	crabs and barnacles	169.21	*Charybdis japonica*	Midway virus (16.3%)
Wuchang Cockroach Virus 3	2203	chuvirus	cockroaches	440.14	*Blattella germanica*	Midway virus (16.3%)
Wuhan Louse Fly Virus 6	2182	chuvirus	insect mix 2	4.12	*unidentified Hippoboscidae*	Midway virus (16.4%)
Wuhan Louse Fly Virus 7	2174	chuvirus	insect mix 2	99.83	*unidentified Hippoboscidae*	Midway virus (17.2%)
Wuhan Mosquito Virus 8	2159	chuvirus	mosquito hubei	300.33	*Culex tritaeniorhynchus*, *C. quinquefasciatus*, *Anopheles sinensis*, *Armigeres subalbatus*	Midway virus (16.7%)
Wuhan Tick Virus 2	2189	chuvirus	ticks	154.46	*Rhipicephalus microplus*	Midway virus (16.7%)
Culex tritaeniorhynchus rhabdovirus	2142	Culex tritaeniorhynchus rhabdovirus	mosquito hubei	3517.32	*Culex tritaeniorhynchus*, *C. quinquefasciatus*, *Anopheles sinensis*, *Armigeres subalbatus*, *Aedes sp*	Isfahan virus (38.5%)
Wuhan Insect virus 4	2105	cytorhabdovirus	insect mix 4	94.92	*Hyalopterus pruni OR Aphelinus sp*	Lettuce necrotic yellows virus (40.6%)
Wuhan Insect virus 5	2098	cytorhabdovirus	insect mix 4	622.97	*Hyalopterus pruni OR Aphelinus sp*	Persimmon virus A (47.9%)
Wuhan Insect virus 6	2079	cytorhabdovirus	insect mix 4	991.99	*Hyalopterus pruni OR Aphelinus sp*	Persimmon virus A (45.2)
Wuhan Louse Fly Virus 5	2123	Kolente virus like	insect mix 2	98.92	*unidentified Hippoboscidae*	Kolente virus (54.5%)
Yongjia Tick Virus 2	2113	Nishimuro virus like	ticks	13.14	*Haemaphysalis hystricis*	Nishimuro virus (54.2%)
Shayang Fly Virus 2	2170	sigmavirus like	true flies	36.83	*Musca domestica*, *Chrysomya megacephala*	Isfahan virus (44.1%)
Wuhan Fly Virus 2	2134	sigmavirus like	true flies	18.37	*Musca domestica*, *Sarcophaga sp*	Vesicular stomatitis Indiana virus (43.4%)
Wuhan House Fly Virus 1	2098	sigmavirus like	true flies	31.04	*Musca domestica*	Isfahan virus (42.8%)
Wuhan Louse Fly Virus 10	2146	sigmavirus like	insect mix 2	235.94	unidentified *Hippoboscidae*	Drosophila melanogaster sigmavirus (51.2%)
Wuhan Louse Fly Virus 8	2145	sigmavirus like	insect mix 2	292.11	unidentified *Hippoboscidae*	Drosophila melanogaster sigmavirus (50.6%)
Wuhan Louse Fly Virus 9	2145	sigmavirus like	insect mix 2	69.37	unidentified *Hippoboscidae*	Drosophila melanogaster sigmavirus (51.4%)
Bole Tick Virus 2	2171	unclassified dimarhabdovirus 1	ticks	38.19	*Hyalomma asiaticum*	Isfahan virus (38.1%)
Huangpi Tick Virus 3	2193	unclassified dimarhabdovirus 1	ticks	15.81	*Haemaphysalis doenitzi*	Eel virus European X (40%)
Tacheng Tick Virus 3	2182	unclassified dimarhabdovirus 1	ticks	96.30	*Dermacentor marginatus*	Eel virus European X (39.8%)
Taishun Tick Virus	2226	unclassified dimarhabdovirus 1	ticks	24.56	*Haemaphysalis hystricis*	Vesicular stomatitis Indiana virus (36.6%)
Wuhan Tick Virus 1	2191	unclassified dimarhabdovirus 1	ticks	119.92	*Rhipicephalus microplus*	Eel virus European X (38.3%)
Wuhan Insect virus 7	2120	unclassified dimarhabdovirus 2	insect mix 4	241.7	*Hyalopterus pruni OR Aphelinus sp*	Isfahan virus (42.6%)
Lishi Spider Virus 2	2201	unclassified mononegavirus 1	spiders	5.57	*unidentified Araneae*	Maize fine streak virus (19.6%)
Sanxia Water Strider Virus 4	2108	unclassified mononegavirus 1	water striders	4767.82	*unidentified Gerridae*	Orchid fleck virus (20.5%)
Tacheng Tick Virus 6	2068	unclassified mononegavirus 1	ticks	17.92	*Argas miniatus*	Maize mosaic virus (20.6%)
Shuangao Fly Virus 2	1966	unclassified mononegavirus 2	insect mix 1	25.94	*Psychoda alternata*	Midway virus (21.3%)
Xincheng Mosquito Virus	2026	unclassified mononegavirus 2	mosquito hubei	400.12	*Anopheles sinensis*	Midway virus (19.2%)
Wenzhou Crab Virus 1	1807	unclassified mononegavirus 3	crabs and barnacles	382.29	*Capitulum mitella*, *Charybdis japonica*, *Charybdis lucifera*	Midway virus (22.2%)
Tacheng Tick Virus 7	2215	unclassified rhabdovirus 1	ticks	35.86	*Argas miniatus*	Orchid fleck virus (24.5%)
Jingshan Fly Virus 2	1970	unclassified rhabdovirus 2	true flies	4.43	*Sarcophaga sp*	Maize fine streak virus (23.4%)
Sanxia Water Strider Virus 5	2264	unclassified rhabdovirus 2	water striders	4373.68	unidentified *Gerridae*	Northern cereal mosaic virus (22.6%)
Shayang Fly Virus 3	2231	unclassified rhabdovirus 2	true flies	27.73	*Chrysomya megacephala*, *Atherigona orientalis*	Maize fine streak virus (22.6%)
Shuangao Bedbug Virus 2	2207	unclassified rhabdovirus 2	insect mix 2	16.29	*Cimex hemipterus*	Maize fine streak virus (22.5%)
Shuangao Insect Virus 6	2088	unclassified rhabdovirus 2	insect mix 1	14.37	unidentified *Diptera*, *Abraxas tenuisuffusa*	Potato yellow dwarf virus (21.2%)
Wuhan Ant Virus	2118	unclassified rhabdovirus 2	insect mix 3	169.79	*Camponotus japonicus*	Lettuce necrotic yellows virus (21.4%)
Wuhan Fly Virus 3	2230	unclassified rhabdovirus 2	true flies	6.00	*Musca domestica*, *Sarcophaga sp*	Maize fine streak virus (21.9%)
Wuhan House Fly Virus 2	2233	unclassified rhabdovirus 2	true flies	221.04	*Musca domestica*	Northern cereal mosaic virus (23.4%)
Wuhan Mosquito Virus 9	2260	unclassified rhabdovirus 2	mosquito hubei	56.19	*Culex tritaeniorhynchus*, *C. quinquefasciatus*, *Aedes sp*	Persimmon virus A (23.2%)
Wuhan Louse Fly Virus 11	2110	Vesiculovirus like	insect mix 2	6.11	unidentified *Hippoboscidae*	Vesicular stomatitis Indiana virus (52.9%)

Table 3. Bunya-arenaviridae-related RdRp sequences discovered in this study

Virus Name	Length of RdRp	Classification	Pool	Abundance	Putative arthropod host	Closest relative (aa identity)
Huangpi Tick Virus 1	3914	Nairovirus like	ticks	11.32	*Haemaphysalis doenitzi*	Hazara virus (39.5%)
Tacheng Tick Virus 1	3962	Nairovirus like	ticks	88.91	*Dermacentor marginatus*	Hazara virus (39.6%)
Wenzhou Tick Virus	3967	Nairovirus like	ticks	44.30	*Haemaphysalis hystricis*	Crimean-Congo hemorrhagic fever virus (39.1%)
Shayang Spider Virus 1	4403	Nairovirus like	spiders	90.95	*Neoscona nautica*, *Parasteatoda tepidariorum*, *Plexippus setipes*	Crimean-Congo hemorrhagic fever virus (26.2%)
Xinzhou Spider Virus	4037	Nairovirus like	spiders	3.79	*Neoscona nautica*, *Parasteatoda tepidariorum*	Erve virus (22.9%)
Sanxia Water Strider Virus 1	3936	Nairovirus like	water striders	26483.38	unidentified *Gerridae*	Hazara virus (23.4%)
Wuhan Louse Fly Virus 1	2250	Orthobunyavirus	insect mix 2	67.06	unidentified *Hippoboscoidea*	La Crosse virus (57.8%)
Shuangao Insect Virus 1	2335	Orthobunyavirus like	insect mix 1	7.97	unidentified *Chrysopidae*, *Psychoda alternata*	Khurdun virus (29.1%)
Wuchang Cockroach Virus 1	2125	phasmavirus like	cockroaches	11283.22	*Blattella germanica*	Kigluaik phantom virus (35.9%)
GAQJ01007189	1554	phasmavirus like	database	N/A	*Ostrinia furnacalis*	Kigluaik phantom virus (35.9%)
Shuangao Insect Virus 2	1765	phasmavirus like	insect mix 1	36.32	*Abraxas tenuisuffusa*, *unidentified diptera*	Kigluaik phantom virus (31.9%)
Wuhan Mosquito Virus 1	2095	phasmavirus like	mosquito Hubei, mosquito Zhejiang	3523.08	*Culex tritaeniorhynchus*, *Anopheles sinensis*, *Culex quinquefasciatus*	Kigluaik phantom virus (39.5%)
Wuhan Mosquito Virus 2	2111	phasmavirus like	mosquito Hubei, mosquito Zhejiang	39.66	*Culex tritaeniorhynchus*, *Anopheles sinensis*, *Culex quinquefasciatus*, *Aedes sp*	Kigluaik phantom virus (39.6%)
Huangpi Tick Virus 2	2121	Phlebovirus	N/A	N/A	*Haemaphysalis sp*	Uukuniemi virus (49.3%)
Bole Tick Virus 1	2148	Phlebovirus	ticks	67.86	*Hyalomma asiaticum*	Uukuniemi virus (37.9%)
Changping Tick Virus 1	2194	Phlebovirus	ticks	335.25	*Dermacentor sp*	Uukuniemi virus (39.7%)
Dabieshan Tick Virus	2148	Phlebovirus	ticks	250.62	*Haemaphysalis longicornis*	Uukuniemi virus (39.2%)
Lihan Tick Virus	2151	Phlebovirus	ticks	60.40	*Rhipicephalus microplus*	Uukuniemi virus (38.6%)
Tacheng Tick Virus 2	2189	Phlebovirus	ticks	132.59	*Dermacentor marginatus*	Uukuniemi virus (39.0%)
Yongjia Tick Virus 1	2138	Phlebovirus	ticks	119.49	*Haemaphysalis hystricis*	Uukuniemi virus (40.5%)
GAIX01000059	2151	Phlebovirus like	database	N/A	*Pararge aegeria*	Cumuto virus (24.1%)
GAKZ01048260	1583	Phlebovirus like	database	N/A	*Procotyla fluviatilis*	Cumuto virus (22.8%)
GAQJ01008681	2261	Phlebovirus like	database	N/A	*Ostrinia furnacalis*	Gouleako virus (22.0%)
Shuangao Insect Virus 3	2050	Phlebovirus like	insect mix 1	339.41	*unidentified Chrysopidae*, *unidentified Diptera*	Cumuto virus (23.7%)
Wuhan Louse Fly Virus 2	2327	Phlebovirus like	insect mix 2	3.57	*unidentified Hippoboscoidea*	Uukuniemi virus (25.2%)
Wuhan Insect virus 1	2099	Phlebovirus like	insect mix 3	178.53	*Asellus sp*, *unidentified Nepidae*, *Camponotus japonicus*	Cumuto virus (24.8%)
Huangshi Humpbacked Fly Virus	2009	Phlebovirus like	insect mix 4	13.13	unidentified *Phoridae*	Cumuto virus (18.1%)
Yichang Insect virus	2100	Phlebovirus like	insect mix 4	71.50	*Aulacorthum magnoliae*	Gouleako virus (45.3%)
Wuhan Millipede Virus 1	1854	Phlebovirus like	millipedes and insect mix 3	825.66	*unidentified Polydesmidae*	Cumuto virus (25.3%)
Qingnian Mosquito Virus	2243	Phlebovirus like	mosquito Hubei	17.09	*Culex quinquefasciatus*	Razdan virus (21.0%)
Wutai Mosquito Virus	2185	Phlebovirus like	mosquito Hubei	70.72	*Culex quinquefasciatus*	Rice stripe virus (26.4%)
Xinzhou Mosquito Virus	2022	Phlebovirus like	mosquito Hubei	98.95	*Anopheles sinensis*	Cumuto virus (24.7%)
Zhee Mosquito Virus	2443	Phlebovirus like	mosquito Hubei, mosquito Zhejiang	308.98	*Anopheles sinensis*, *Armigeres subalbatus*	Cumuto virus (22.6%)
Whenzhou Shrimp Virus 1	2051	Phlebovirus like	shrimps	5859.37	*Penaeus monodon*	Uukuniemi virus (32.2%)
Wuhan Spider Virus	2251	Phlebovirus like	spiders	17.71	*Neoscona nautica*, *Parasteatoda tepidariorum*, *Plexippus setipes*	Uukuniemi virus (21.7%)
Wuhan Fly Virus 1	2192	Phlebovirus like	true flies	68.58	*Atherigona orientalis*, *Chrysomya megacephala*, *Sarcophaga sp*, *Musca domestica*	Grand Arbaud virus (27.8%)
Wuhan horsefly Virus	3117	Tenuivirus like	horseflies	13.50	unidentified *Tabanidae*	Uukuniemi virus (28.2%)
Jiangxia Mosquito Virus 1	1889	Unclassified segmented virus 1	mosquito Hubei	11.55	*Culex tritaeniorhynchus*	Gouleako virus (16.7%)
Shuangao Bedbug Virus 1	2015	Unclassified segmented virus 2	insect mix 2	12.71	*Cimex hemipterus*	Murrumbidgee virus (16.3%)
Jiangxia Mosquito Virus 2	1860	Unclassified segmented virus 2	mosquito Hubei	2.81	*Culex tritaeniorhynchus*	Hantavirus (18.9%)
Shuangao Mosquito Virus	1996	Unclassified segmented virus 2	mosquito Zhejiang	11.67	*Armigeres subalbatus*	Hantavirus (18.7%)
Whenzhou Shrimp Virus 2	2241	Unclassified segmented virus 3	shrimps	3824.55	*Penaeus monodon*, *Exopalaemon carinicauda*	La Crosse virus (19.0%)
Shayang Spider Virus 2	2165	Unclassified segmented virus 4	spiders	12.75	*Neoscona nautica*, *Pirata sp*, *Parasteatoda tepidariorum*, unidentified *Araneae*	Akabane virus (16.6%)
Wuhan Insect virus 2	2377	Unclassified segmented virus 5	insect mix 4	223.06	*Hyalopterus pruni OR Aphelinus sp*	Kigluaik phantom virus (19.2%)
Sanxia Water Strider Virus 2	2349	Unclassified segmented virus 5	water striders	707.09	unidentified *Gerridae*	Kigluaik phantom virus (19.8%)
Wuhan Millipede Virus 2	3709	Unclassified segmented virus 6	millipedes	1513.41	unidentified *Polydesmidae*	Dugbe virus (17.2%)
Wuhan Insect virus 3	2231	Unclassified segmented virus 7	insect mix 3	3.50	*Asellus sp*	Herbert virus (17.2%)

649 **Table 4**. Orthomyxoviridae-related RdRp sequences discovered in this study

Virus Name	Length of RdRp	Classification	Pool	Abundance	Putative arthropod host	Closest relative (aa identity)
Jingshan Fly Virus 1	795	Quaranjavirus	true flies	21.93	*Atherigona orientalis*, *Chrysomya megacephala*, *Sarcophaga sp*, *Musca domestica*	Johnston Atoll virus (36.9%)
Jiujie Fly Virus	653	Quaranjavirus	horseflies	10.30	unidentified *Tabanidae*	Johnston Atoll virus (39.7%)
Sanxia Water Strider Virus 3	789	Quaranjavirus	water striders	1101.03	unidentified *Gerridae*	Johnston Atoll virus (36.7%)
Shayang Spider Virus 3	768	Quaranjavirus	spiders	1.95	*Neoscona nautica*	Johnston Atoll virus (38.5%)
Shuangao Insect Virus 4	793	Quaranjavirus	insect mix1	59.90	unidentified *Diptera*, unidentified *Stratiomyidae*	Johnston Atoll virus (36.9%)
Wuhan Louse Fly Virus 3	784	Quaranjavirus	insect mix2	500.77	unidentified *Hippoboscoidea*	Johnston Atoll virus (37.7%)
Wuhan Louse Fly Virus 4	783	Quaranjavirus	insect mix2	96.80	unidentified *Hippoboscoidea*	Johnston Atoll virus (38.2%)
Wuhan Mosquito Virus 3	801	Quaranjavirus	mosquito Hubei	40.07	*Culex tritaeniorhynchus*, *Culex quinquefasciatus*, *Armigeres subalbatus*	Johnston Atoll virus (35.6%)
Wuhan Mosquito Virus 4	792	Quaranjavirus	mosquito Hubei	86.21	*Culex tritaeniorhynchus*, *Culex quinquefasciatus*, *Armigeres subalbatus*	Johnston Atoll virus (34.8%)
Wuhan Mosquito Virus 5	806	Quaranjavirus	mosquito Hubei	75.05	*Culex tritaeniorhynchus*, *Culex quinquefasciatus*, *Armigeres subalbatus*	Johnston Atoll virus (35.5%)
Wuhan Mosquito Virus 6	800	Quaranjavirus	mosquito Hubei	56.30	*Culex quinquefasciatus*	Johnston Atoll virus (34.2%)
Wuhan Mosquito Virus 7	779	Quaranjavirus	mosquito Hubei	20.74	*Anopheles sinensis*, *Culex quinquefasciatus*	Johnston Atoll virus (34.1%)
Wuhan Mothfly Virus	710	Quaranjavirus	insect mix4	14.47	*Psychoda alternata*	Johnston Atoll virus (39.7%)
Wuchang Cockroach Virus 2	671	Unclassified orthomyxovirus 1	cockroaches	4.01	*Blattella germanica*	Influenza C virus (27.0%)

650

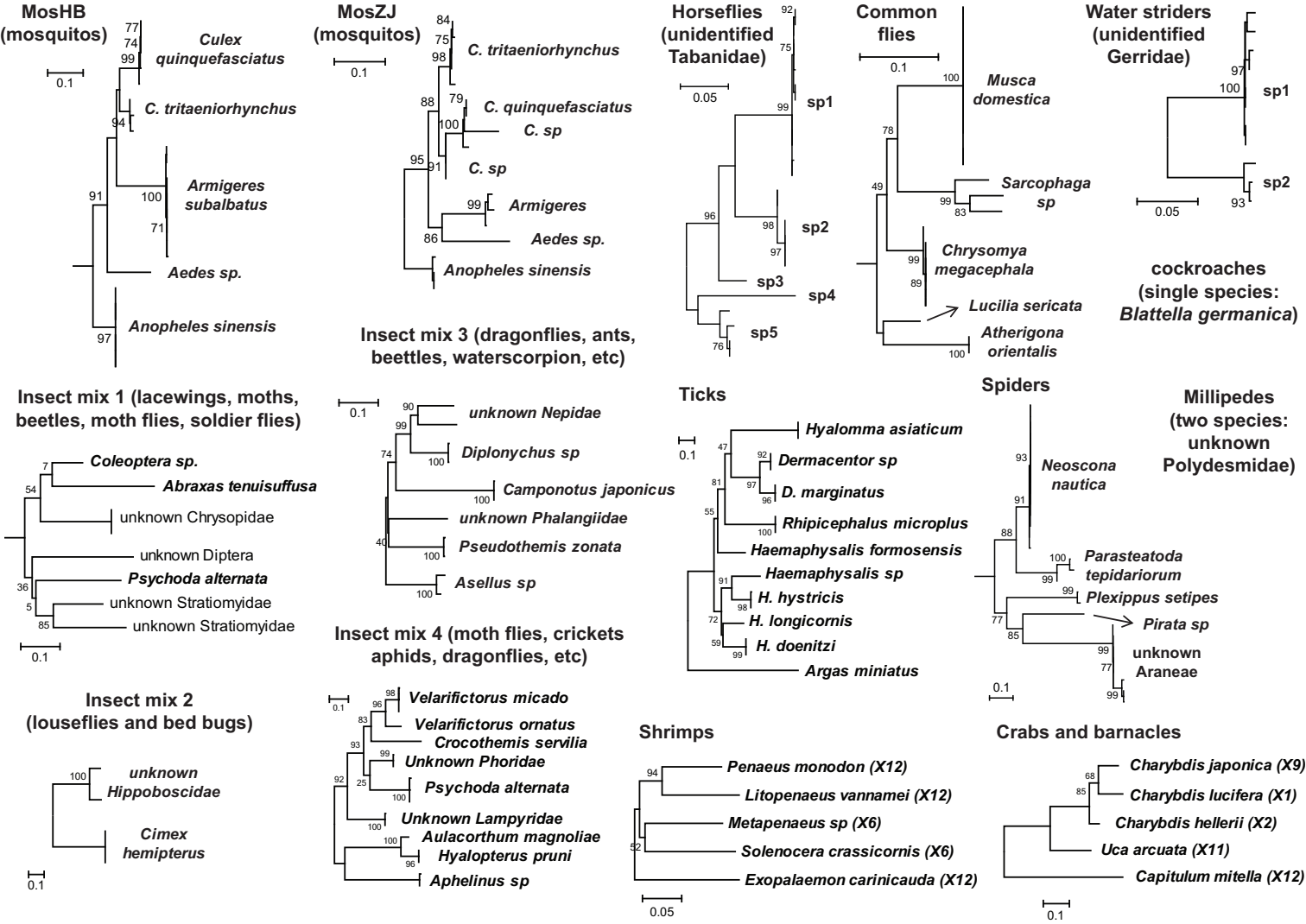
651

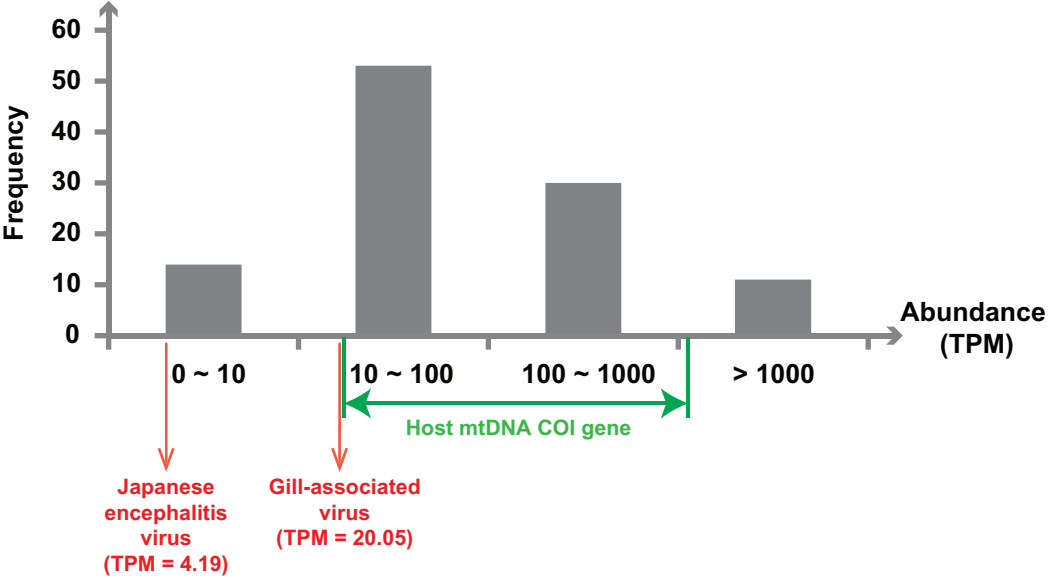
652

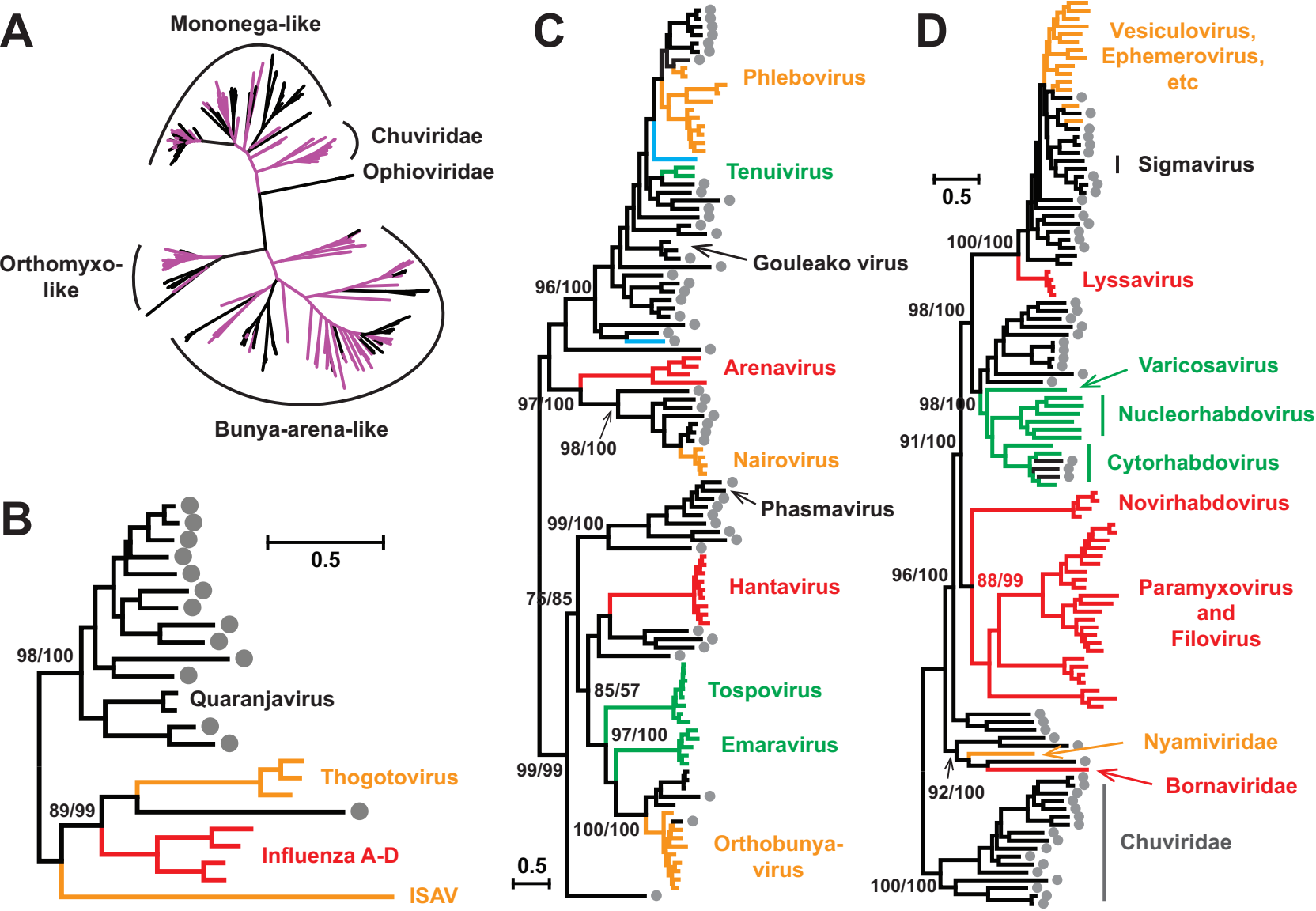
653 **Table 5.** Summary of Endogenous Virus Elements (EVEs) determined here

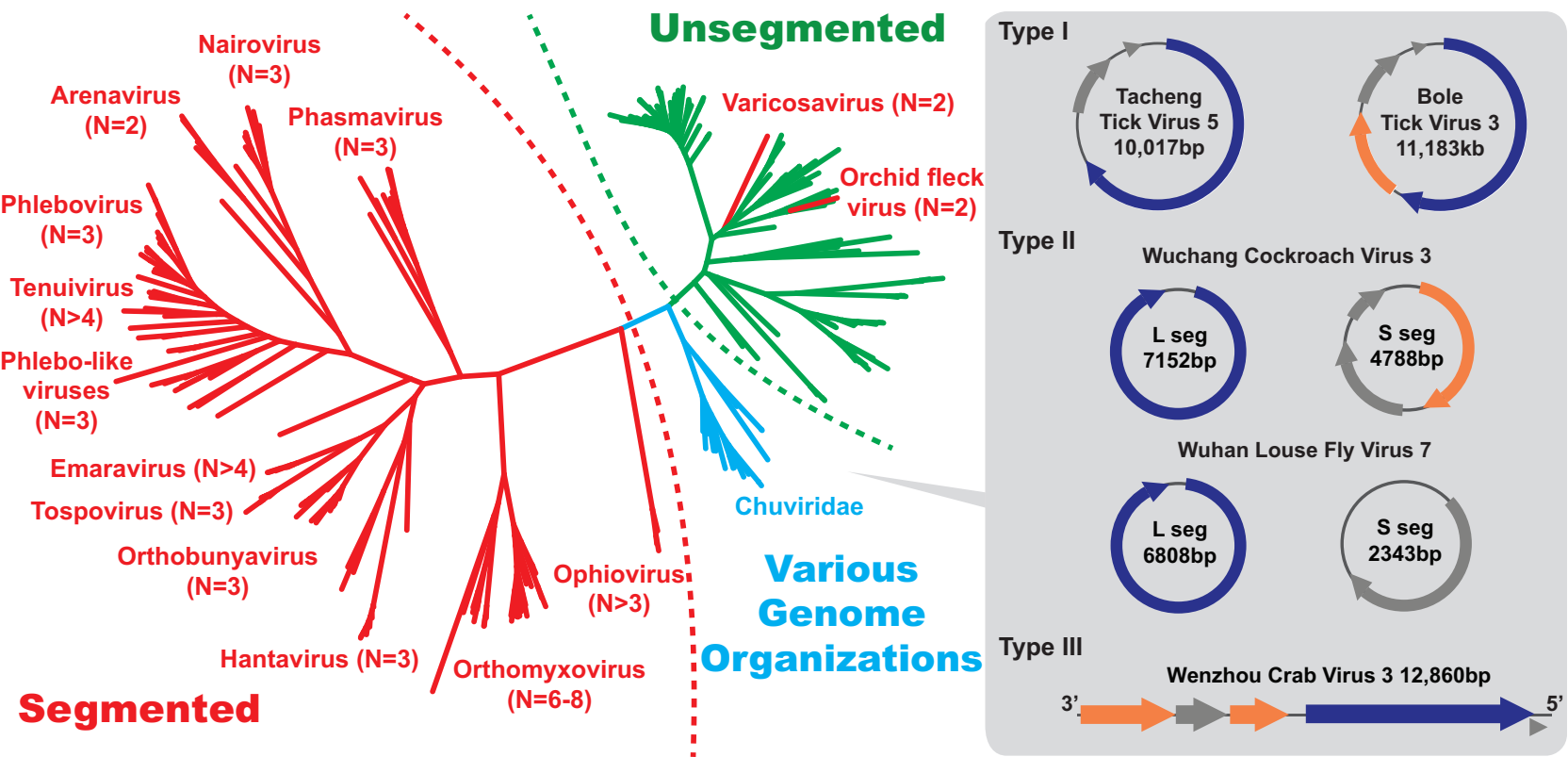
Host classification	Host name	Virus classification	Gene(s) present
Chelicerata	*Ixodes scapularis*	Chuvirus	G, N
		Dimarhabdovirus	RdRp, N
		Nairovirus like	N
		Phlebovirus	RdRp, N
		Quaranjavirus	RdRp
	Tetranychus urticae	Dimarhabdovirus	N
Crustacea	*Daphnia pulex*	Phlebovirus like	RdRp
	Eurytemora affinis	Chuvirus	G
		Dimarhabdovirus	RdRp, N
	Hyalella azteca	Chuvirus	G, N
		Unclassified mononegavirus 3	RdRp, N
	Lepeophtheirus salmonis	Phlebovirus like	N, G
Insecta: Coleoptera	*Dendroctonus ponderosae*	Chuvirus	G
		Phasmavirus	G, N
	Tribolium castaneum	Chuvirus	G
Insecta: Diptera	*Aedes aegypti*	Chuvirus	RdRp
		Dimarhabdovirus	RdRp, N
		Phasmavirus	G
		Phlebovirus like	N
		Quaranjavirus	RdRp
	Anopheles spp.	Chuvirus	G
		Dimarhabdovirus	RdRp, N
		Phasmavirus	G, N
		Phlebovirus like	N
		Quaranjavirus	RdRp
	Culex quinquefasciatus	Chuvirus	G, N
		Dimarhabdovirus	N
	Drosophila spp.	Dimarhabdovirus	RdRp, N
		Phasmavirus	N
		Unclassified rhabdovirus 2	RdRp, N
Insecta: Isoptera	*Zootermopsis nevadensis*	Chuvirus	N
Insecta: Hemiptera	*Acyrthosiphon pisum*	Chuvirus	G, N
		Dimarhabdovirus	N
		Phlebovirus like	N
		Quaranjavirus	RdRp
		Unclassified mononegavirus 1	RdRp, N
	Rhodnius prolixus	Chuvirus	G
		Phasmavirus	G
Insecta: Hymenoptera	*Atta cephalotes*	Unclassified mononegavirus 2	RdRp
	Acromyrmex echinatior	Chuvirus	G
		Unclassified mononegavirus 2	RdRp
	Camponotus floridanus	Chuvirus	G
		Unclassified mononegavirus 1	N
		Unclassified mononegavirus 3	RdRp
		Unclassified rhabdovirus 2	RdRp
	Harpegnathos saltator	Chuvirus	G
	Linepithema humile	Chuvirus	G
	Nasonia spp.	Chuvirus	G
	Pogonomyrmex barbatus	Chuvirus	G
	Solenopsis invicta	Chuvirus	G
		Unclassified mononegavirus 1	N
		Unclassified mononegavirus 3	RdRp, N
Insecta: Lepidoptera	*Bombyx mori*	Chuvirus	RdRp, G
		Quaranjavirus	RdRp
		Unclassified rhabdovirus 2	RdRp
	Melitaea cinxia	Dimarhabdovirus	N
		Quaranjavirus	RdRp
	Plutella xylostella	Dimarhabdovirus	N, G
	Spodoptera frugiperda	Phlebovirus like	G
Myriapoda	*Strigamia maritima*	Chuvirus	N
		Phlebovirus like	G

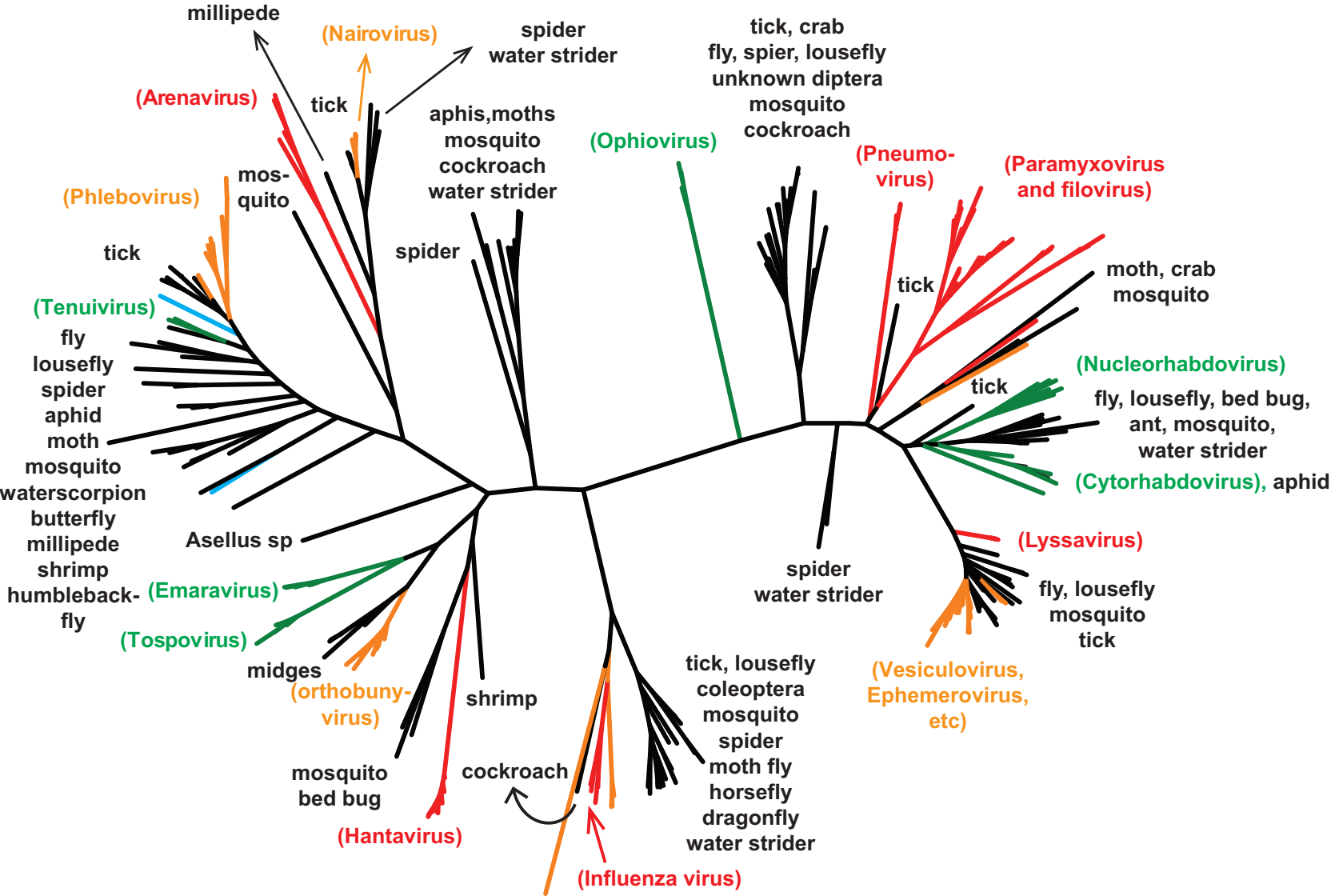
654

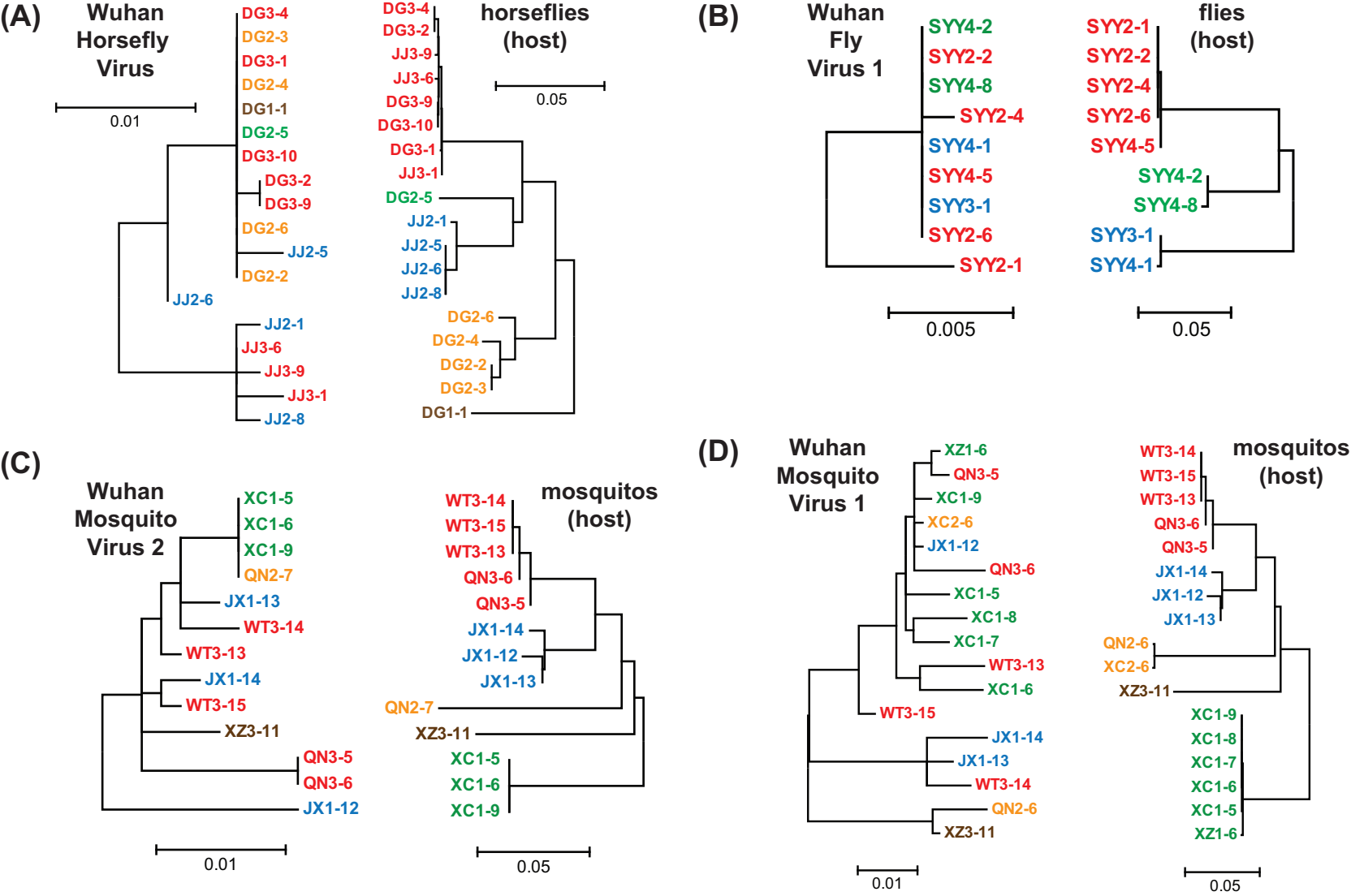

MosHB
(mosquitos)
Culex quinquefasciatus
C. tritaeniorhynchus
Armigeres subalbatus
Aedes sp.
Anopheles sinensis
MosZJ
(mosquitos)
C. tritaeniorhynchus
C. quinquefasciatus
C. sp
C. sp
Armigeres
Aedes sp.
Anopheles sinensis
Horseflies
(unidentified
Tabanidae)
sp1
sp2
sp3
sp4
sp5
Common
flies
Musca domestica
Sarcophaga sp
Chrysomya megacephala
Lucilia sericata
Atherigona orientalis
Water striders
(unidentified
Gerridae)
sp1
sp2
cockroaches
(single species:
Blattella germanica)
Insect mix 1 (lacewings, moths,
beetles, moth flies, soldier flies)
Coleoptera sp.
Abraxas tenuisuffusa
unknown Chrysopidae
unknown Diptera
Psychoda alternata
unknown Stratiomyidae
unknown Stratiomyidae
Insect mix 3 (dragonflies, ants,
beettles, waterscorpion, etc)
unknown Nepidae
Diplonychus sp
Camponotus japonicus
unknown Phalangiidae
Pseudothemis zonata
Asellus sp
Ticks
Hyalomma asiaticum
Dermacentor sp
D. marginatus
Rhipicephalus microplus
Haemaphysalis formosensis
Haemaphysalis sp
H. hystricis
H. longicornis
H. doenitzi
Argas miniatus
Spiders
Neoscona nautica
Parasteatoda tepidariorum
Plexippus setipes
Pirata sp
unknown
Araneae
Millipedes
(two species:
unknown
Polydesmidae)
Insect mix 2
(louseflies and bed bugs)
unknown
Hippoboscidae
Cimex hemipterus
Insect mix 4 (moth flies, crickets
aphids, dragonflies, etc)
Velarifictorus micado
Velarifictorus ornatus
Crocothemis servilia
Unknown Phoridae
Psychoda alternata
Unknown Lampyridae
Aulacorthum magnoliae
Hyalopterus pruni
Aphelinus sp
Shrimps
Penaeus monodon (X12)
Litopenaeus vannamei (X12)
Metapenaeus sp (X6)
Solenocera crassicornis (X6)
Exopalaemon carinicauda (X12)
Crabs and barnacles
Charybdis japonica (X9)
Charybdis lucifera (X1)
Charybdis hellerii (X2)
Uca arcuata (X11)
Capitulum mitella (X12)

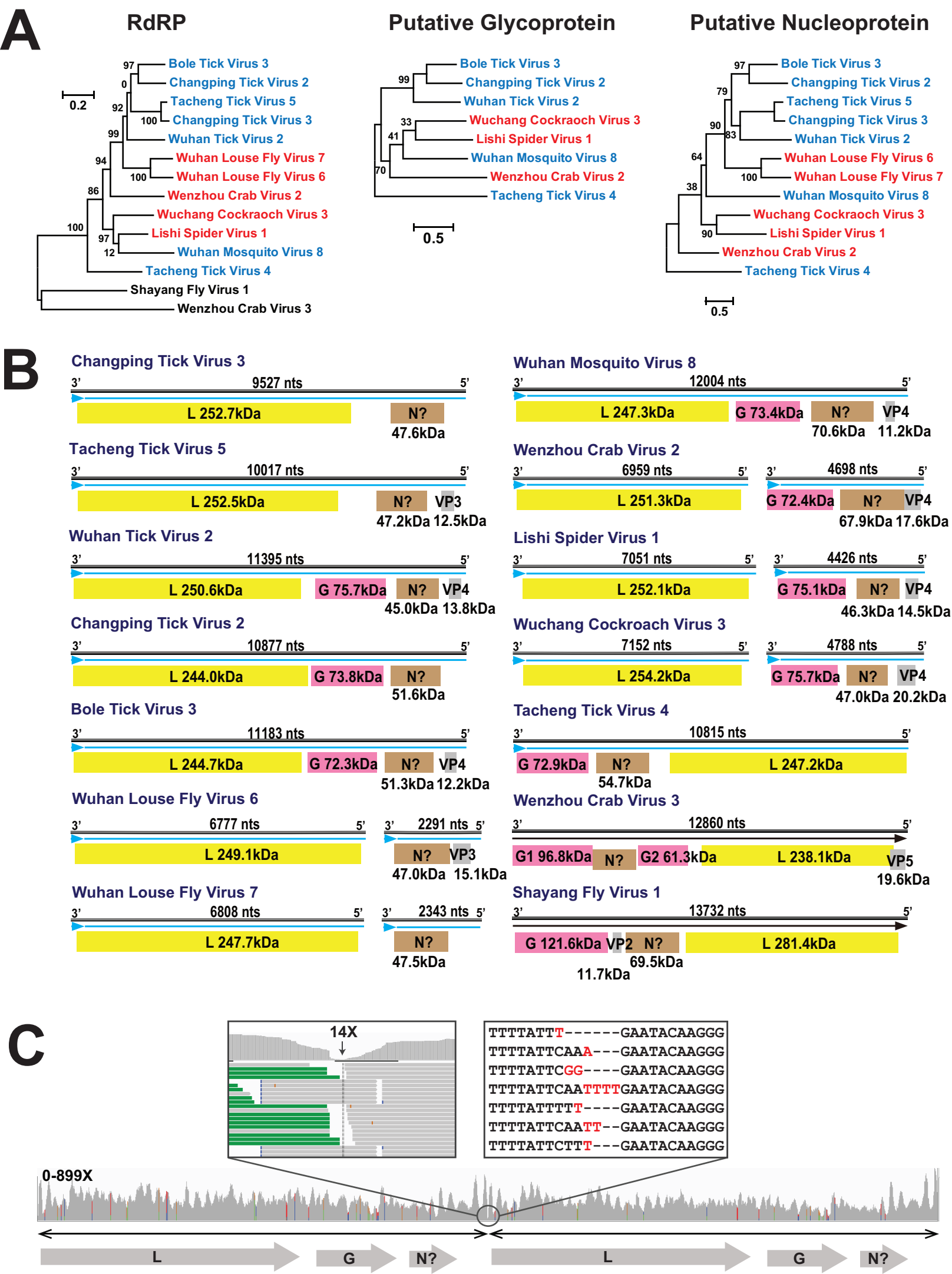

Frequency
60
50
40
30
20
10
0
Abundance
(TPM)
0 ~ 10
10 ~ 100
100 ~ 1000
> 1000
Host mtDNA COI gene
Japanese encephalitis virus (TPM = 4.19)
Gill-associated virus (TPM = 20.05)

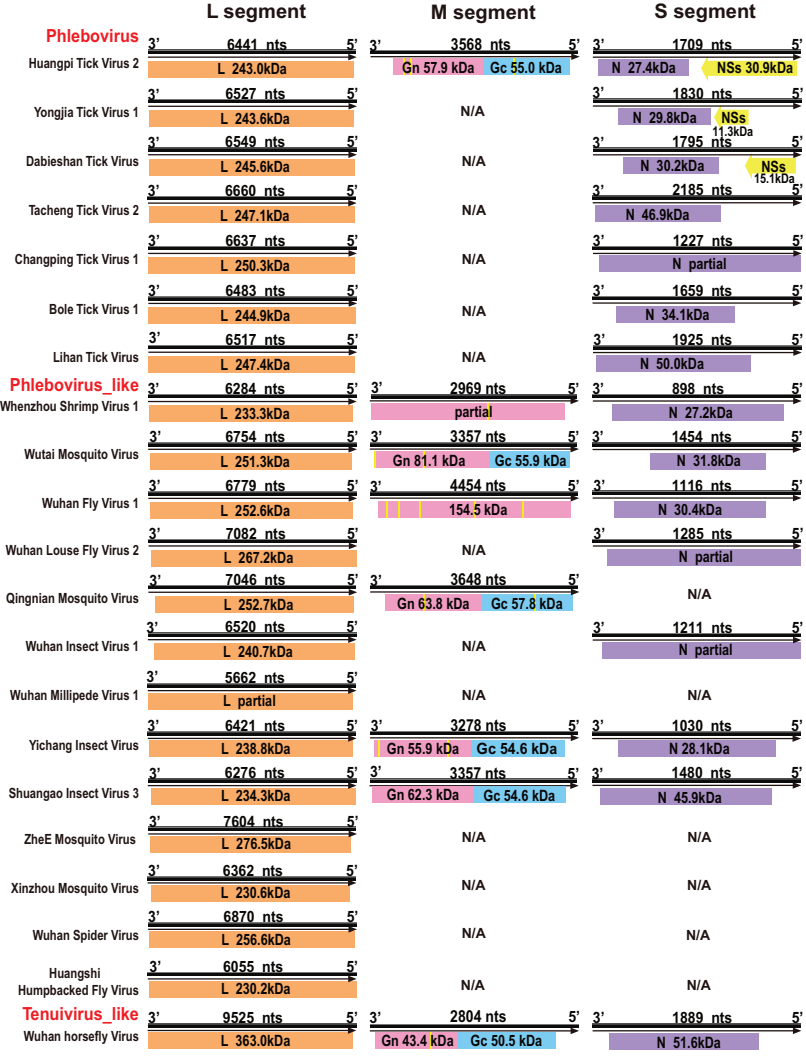

A
Mononega-like
Chuviridae
Ophioviridae
Orthomyxo-
like
Bunya-arena-like
B
0.5
98/100
Quaranjavirus
Thogotovirus
89/99
Influenza A-D
ISAV
C
Phlebovirus
Tenuivirus
Gouleako virus
96/100
Arenavirus
97/100
98/100
Nairovirus
Phasmavirus
99/100
Hantavirus
75/85
85/57
Tospovirus
97/100
Emaravirus
99/99
100/100
Orthobunya-
virus
0.5
D
Vesiculovirus,
Ephemerovirus,
etc
Sigmavirus
0.5
100/100
Lyssavirus
98/100
Varicosavirus
98/100
Nucleorhabdovirus
91/100
Cytorhabdovirus
Novirhabdovirus
96/100
88/99
Paramyxovirus
and
Filovirus
Nyamiviridae
92/100
Bornaviridae
Chuviridae
100/100


Unsegmented
Nairovirus (N=3)
Arenavirus (N=2)
Phasmavirus (N=3)
Varicosavirus (N=2)
Orchid fleck virus (N=2)
Phlebovirus (N=3)
Tenuivirus (N>4)
Phlebo-like viruses (N=3)
Emaravirus (N>4)
Tospovirus (N=3)
Orthobunyavirus (N=3)
Chuviridae
Ophiovirus (N>3)
Hantavirus (N=3)
Orthomyxovirus (N=6-8)
Segmented
Various Genome Organizations
Type I
Tacheng Tick Virus 5 10,017bp
Bole Tick Virus 3 11,183kb
Type II
Wuchang Cockroach Virus 3
L seg 7152bp
S seg 4788bp
Wuhan Louse Fly Virus 7
L seg 6808bp
S seg 2343bp
Type III
Wenzhou Crab Virus 3 12,860bp
3'
5'


millipede
(Nairovirus)
spider
water strider
(Arenavirus)
tick
aphis,moths
mosquito
cockroach
water strider
tick, crab
fly, spier, lousefly
unknown diptera
mosquito
cockroach
(Ophiovirus)
(Pneumo-
virus)
(Paramyxovirus
and filovirus)
(Phlebovirus)
mos-
quito
tick
spider
tick
moth, crab
mosquito
(Tenuivirus)
fly
lousefly
spider
aphid
moth
mosquito
waterscorpion
butterfly
millipede
shrimp
humbleback-
fly
(Nucleorhabdovirus)
tick
fly, lousefly, bed bug,
ant, mosquito,
water strider
(Cytorhabdovirus), aphid
Asellus sp
(Lyssavirus)
spider
water strider
fly, lousefly
mosquito
tick
(Emaravirus)
(Tospovirus)
midges
(orthobuny-
virus)
shrimp
tick, lousefly
coleoptera
mosquito
spider
moth fly
horsefly
dragonfly
water strider
(Vesiculovirus,
Ephemerovirus,
etc)
mosquito
bed bug
cockroach
(Hantavirus)
(Influenza virus)

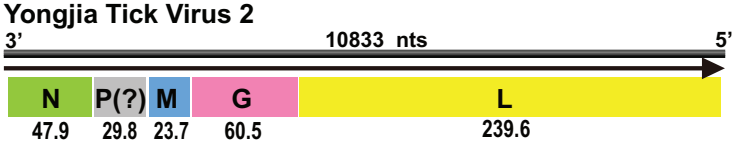

(A)
Wuhan Horsefly Virus
DG3-4
DG2-3
DG3-1
DG2-4
DG1-1
DG2-5
DG3-10
DG3-2
DG3-9
DG2-6
JJ2-5
DG2-2
JJ2-6
JJ2-1
JJ3-6
JJ3-9
JJ3-1
JJ2-8
0.01
horseflies (host)
DG3-4
DG3-2
JJ3-9
JJ3-6
DG3-9
DG3-10
DG3-1
JJ3-1
DG2-5
JJ2-1
JJ2-5
JJ2-6
JJ2-8
DG2-6
DG2-4
DG2-2
DG2-3
DG1-1
0.05
(B)
Wuhan Fly Virus 1
SYY4-2
SYY2-2
SYY4-8
SYY2-4
SYY4-1
SYY4-5
SYY3-1
SYY2-6
SYY2-1
0.005
flies (host)
SYY2-1
SYY2-2
SYY2-4
SYY2-6
SYY4-5
SYY4-2
SYY4-8
SYY3-1
SYY4-1
0.05
(C)
Wuhan Mosquito Virus 2
XC1-5
XC1-6
XC1-9
QN2-7
JX1-13
WT3-14
WT3-13
JX1-14
WT3-15
XZ3-11
QN3-5
QN3-6
JX1-12
0.01
mosquitos (host)
WT3-14
WT3-15
WT3-13
QN3-6
QN3-5
JX1-14
JX1-12
JX1-13
QN2-7
XZ3-11
XC1-5
XC1-6
XC1-9
0.05
(D)
Wuhan Mosquito Virus 1
XZ1-6
QN3-5
XC1-9
XC2-6
JX1-12
QN3-6
XC1-5
XC1-8
XC1-7
WT3-13
XC1-6
WT3-15
JX1-14
JX1-13
WT3-14
QN2-6
XZ3-11
0.01
mosquitos (host)
WT3-14
WT3-15
WT3-13
QN3-6
QN3-5
JX1-14
JX1-12
JX1-13
QN2-6
XC2-6
XZ3-11
XC1-9
XC1-8
XC1-7
XC1-6
XC1-5
XZ1-6
0.05


A
RdRP
Bole Tick Virus 3
Changping Tick Virus 2
Tacheng Tick Virus 5
Changping Tick Virus 3
Wuhan Tick Virus 2
Wuhan Louse Fly Virus 7
Wuhan Louse Fly Virus 6
Wenzhou Crab Virus 2
Wuchang Cockraoch Virus 3
Lishi Spider Virus 1
Wuhan Mosquito Virus 8
Tacheng Tick Virus 4
Shayang Fly Virus 1
Wenzhou Crab Virus 3
0.2
Putative Glycoprotein
Bole Tick Virus 3
Changping Tick Virus 2
Wuhan Tick Virus 2
Wuchang Cockraoch Virus 3
Lishi Spider Virus 1
Wuhan Mosquito Virus 8
Wenzhou Crab Virus 2
Tacheng Tick Virus 4
0.5
Putative Nucleoprotein
Bole Tick Virus 3
Changping Tick Virus 2
Tacheng Tick Virus 5
Changping Tick Virus 3
Wuhan Tick Virus 2
Wuhan Louse Fly Virus 6
Wuhan Louse Fly Virus 7
Wuhan Mosquito Virus 8
Wuchang Cockraoch Virus 3
Lishi Spider Virus 1
Wenzhou Crab Virus 2
Tacheng Tick Virus 4
0.5
B
Changping Tick Virus 3
9527 nts
L 252.7kDa
N?
47.6kDa
Tacheng Tick Virus 5
10017 nts
L 252.5kDa
N?
VP3
47.2kDa 12.5kDa
Wuhan Tick Virus 2
11395 nts
L 250.6kDa
G 75.7kDa
N?
VP4
45.0kDa 13.8kDa
Changping Tick Virus 2
10877 nts
L 244.0kDa
G 73.8kDa
N?
51.6kDa
Bole Tick Virus 3
11183 nts
L 244.7kDa
G 72.3kDa
N?
VP4
51.3kDa 12.2kDa
Wuhan Louse Fly Virus 6
6777 nts
L 249.1kDa
2291 nts
N?
VP3
47.0kDa 15.1kDa
Wuhan Louse Fly Virus 7
6808 nts
L 247.7kDa
2343 nts
N?
47.5kDa
Wuhan Mosquito Virus 8
12004 nts
L 247.3kDa
G 73.4kDa
N?
VP4
70.6kDa 11.2kDa
Wenzhou Crab Virus 2
6959 nts
L 251.3kDa
4698 nts
G 72.4kDa
N?
VP4
67.9kDa 17.6kDa
Lishi Spider Virus 1
7051 nts
L 252.1kDa
4426 nts
G 75.1kDa
N?
VP4
46.3kDa 14.5kDa
Wuchang Cockroach Virus 3
7152 nts
L 254.2kDa
4788 nts
G 75.7kDa
N?
VP4
47.0kDa 20.2kDa
Tacheng Tick Virus 4
10815 nts
G 72.9kDa
N?
54.7kDa
L 247.2kDa
Wenzhou Crab Virus 3
12860 nts
G1 96.8kDa
N?
G2 61.3kDa
L 238.1kDa
VP5
19.6kDa
Shayang Fly Virus 1
13732 nts
G 121.6kDa
VP2
N?
69.5kDa
11.7kDa
L 281.4kDa
C
14X
TTTTATTT------GAATACAAGGG
TTTTATTCAAA---GAATACAAGGG
TTTTATTCGG----GAATACAAGGG
TTTTATTCAATTTTGAATACAAGGG
TTTTATTTTT----GAATACAAGGG
TTTTATTCAATT--GAATACAAGGG
TTTTATTCTTT---GAATACAAGGG
0-899X
L
G
N?
L
G
N?


L segment
M segment
S segment
Phlebovirus
Huangpi Tick Virus 2
L 6441 nts
L 243.0kDa
M 3568 nts
Gn 57.9 kDa
Gc 55.0 kDa
S 1709 nts
N 27.4kDa
NSs 30.9kDa
Yongjia Tick Virus 1
L 6527 nts
L 243.6kDa
N/A
S 1830 nts
N 29.8kDa
NSs 11.3kDa
Dabieshan Tick Virus
L 6549 nts
L 245.6kDa
N/A
S 1795 nts
N 30.2kDa
NSs 15.1kDa
Tacheng Tick Virus 2
L 6660 nts
L 247.1kDa
N/A
S 2185 nts
N 46.9kDa
Changping Tick Virus 1
L 6637 nts
L 250.3kDa
N/A
S 1227 nts
N partial
Bole Tick Virus 1
L 6483 nts
L 244.9kDa
N/A
S 1659 nts
N 34.1kDa
Lihan Tick Virus
L 6517 nts
L 247.4kDa
N/A
S 1925 nts
N 50.0kDa
Phlebovirus_like
Whenzhou Shrimp Virus 1
L 6284 nts
L 233.3kDa
M 2969 nts
partial
S 898 nts
N 27.2kDa
Wutai Mosquito Virus
L 6754 nts
L 251.3kDa
M 3357 nts
Gn 81.1 kDa
Gc 55.9 kDa
S 1454 nts
N 31.8kDa
Wuhan Fly Virus 1
L 6779 nts
L 252.6kDa
M 4454 nts
154.5 kDa
S 1116 nts
N 30.4kDa
Wuhan Louse Fly Virus 2
L 7082 nts
L 267.2kDa
N/A
S 1285 nts
N partial
Qingnian Mosquito Virus
L 7046 nts
L 252.7kDa
M 3648 nts
Gn 63.8 kDa
Gc 57.8 kDa
N/A
Wuhan Insect Virus 1
L 6520 nts
L 240.7kDa
N/A
S 1211 nts
N partial
Wuhan Millipede Virus 1
L 5662 nts
L partial
N/A
N/A
Yichang Insect Virus
L 6421 nts
L 238.8kDa
M 3278 nts
Gn 55.9 kDa
Gc 54.6 kDa
S 1030 nts
N 28.1kDa
Shuangao Insect Virus 3
L 6276 nts
L 234.3kDa
M 3357 nts
Gn 62.3 kDa
Gc 54.6 kDa
S 1480 nts
N 45.9kDa
ZheE Mosquito Virus
L 7604 nts
L 276.5kDa
N/A
N/A
Xinzhou Mosquito Virus
L 6362 nts
L 230.6kDa
N/A
N/A
Wuhan Spider Virus
L 6870 nts
L 256.6kDa
N/A
N/A
Huangshi Humpbacked Fly Virus
L 6055 nts
L 230.2kDa
N/A
N/A
Tenuivirus_like
Wuhan horsefly Virus
L 9525 nts
L 363.0kDa
M 2804 nts
Gn 43.4 kDa
Gc 50.5 kDa
S 1889 nts
N 51.6kDa
L segment
M segment
S segment
Nairovirus_like
Xinzhou Spider Virus
L 12336 nts
L 464.6kDa
M 3322 nts
Gn 66.3 kDa
Gc partial
S 2241 nts
N 56.6kDa
Sanxia Water Strider virus 1
L 12229 nts
L 450.0kDa
M 4193 nts
Gn 52.2 kDa
Gc 74.9 kDa
S 2171 nts
N 53.9kDa
Shayang Spider Virus 1
L 13486 nts
L 506.5kDa
M 4805 nts
Gn 60.2 kDa
Gc 100.9 kDa
S 1940 nts
N 64.2kDa
Huangpi Tick Virus 1
L 11823 nts
L 444.7kDa
M 4675 nts
Gn 66.1 kDa
Gc 77.9 kDa
S 1969 nts
N 53.9kDa
Wenzhou Tick Virus
L 12073 nts
L 441.4kDa
M 4428 nts
Gn 72.0 kDa
Gc 73.7 kDa
S 1785 nts
N 54.7kDa
Tacheng Tick Virus 1
L 12184 nts
L 446.9kDa
M 5147 nts
Gn 73.2 kDa
Gc 76.1 kDa
S 1840 nts
N 54.6kDa
Phasmavirus_like
Wuhan Mosquito Virus 2
L 6574 nts
L 242.4kDa
M 2073 nts
Gn
Gc 52.8 kDa
S 2208 nts
NSs (?)
N 42.9kDa
Wuhan Mosquito Virus 1
L 6474 nts
L 238.4kDa
M 2217 nts
partial
Gn
Gc 53.2 kDa
S 2111 nts
NSs(?) 14.1kDa
N 38.2kDa
VP3(?) 12.3kDa
Wuchang Cockraoch Virus 1
L 6493 nts
L 246.1kDa
M 2636 nts
partial
Gn 35.1 kDa
Gc 53.7 kDa
S 1808 nts
NSs (?) 13.9kDa
N 48.3kDa
Shuangao Insect Virus 2
L 4899 nts
L partial
M 2255 nts
Gn 23.1 kDa
Gc 51.7 kDa
S 1959 nts
NSs (?) 15.4kDa
N 54.7kDa
VP3 (?) 18.4kDa
S 1148 nts
NSs (?)
N partial
Orthobunyavirus
Wuhan Louse Fly Virus 1
L 6787 nts
L 262.4kDa
M 2939 nts
Gn
NSm 15.9kDa
partial
Gc partial
S 959 nts
NSs 9.6kDa
N 27.1kDa
Orthobunyavirus_like
Shuangao Insect Virus 1
L 7325 nts
L 264.8kDa
M 4753 nts
Gn
NSm 27.4kDa
Gc 111.4 kDa
S 1233 nts
N 30.0kDa
Unclassified segmented virus 1
Jiangxia Mosquito Virus 1
L 5754 nts
L partial
M 2957 nts
Gn partial
Gc partial
N/A
Unclassified segmented virus 2
Jiangxia Mosquito Virus 2
L 5956 nts
L 217.0kDa
M 2312 nts
Gn
partial
Gc partial
N/A
Shuangao Bedbug Virus 1
L 6085 nts
L 229.7kDa
M 3302 nts
Gn 49.0 kDa
Gc 67.7 kDa
N/A
Shuangao Mosquito Virus
L 6553 nts
L partial
N/A
N/A
Unclassified segmented virus 3
Whenzhou Shrimp Virus 2
L 6873 nts
L 258.0kDa
M 2466 nts
Gn
partial
Gc 56.3 kDa
N/A
Unclassified segmented virus 4
Shayang Spider Virus 2
L 6646 nts
L 249.4kDa
N/A
N/A
Unclassified segmented virus 5
Wuhan Insect Virus 2
L 7243 nts
L 272.2kDa
M 2852 nts
Gn partial
Gc partial
S 1820 nts
N 53.9kDa
Sanxia Water Strider Virus 2
L 7151 nts
L 271.6kDa
N/A
N/A
Unclassified segmented virus 6
Wuhan Millipede Virus 2
L 11332 nts
L 417.0kDa
N/A
S 1960 nts
N 53.8kDa
Unclassified segmented virus 7
Wuhan Insect Virus 3
L 6743 nts
L 249.0kDa
N/A
S 1867 nts
N 53.8kDa

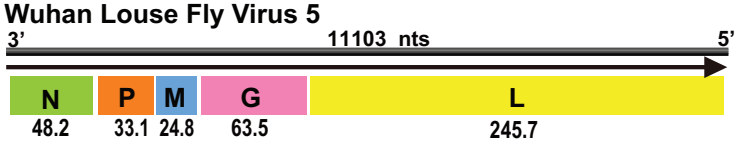
Nishimuro virus like

Yongjia Tick Virus 2



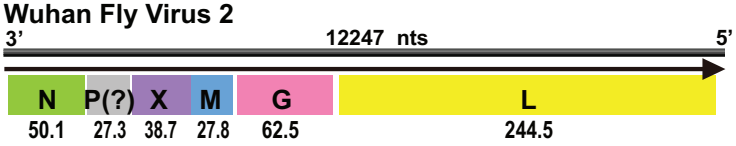
Kolente virus like

Wuhan Louse Fly Virus 5

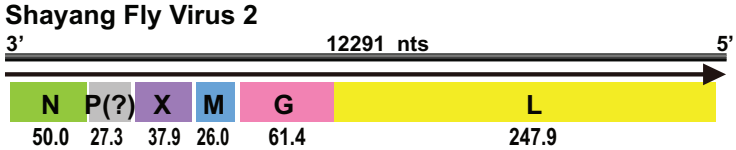


Sigmavirus_like

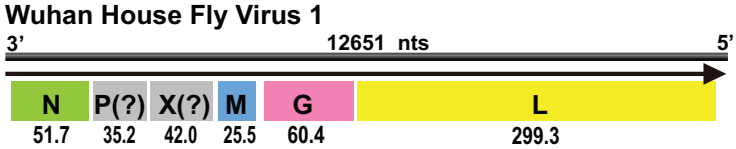
Wuhan Fly Virus 2



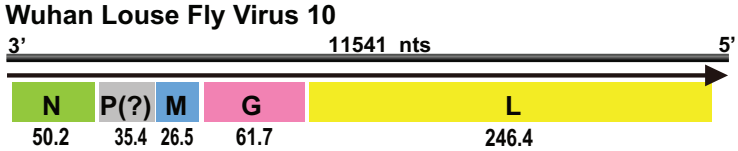
Shayang Fly Virus 2



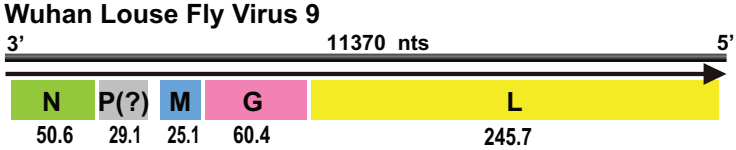
Wuhan House Fly Virus 1



Wuhan Louse Fly Virus 10

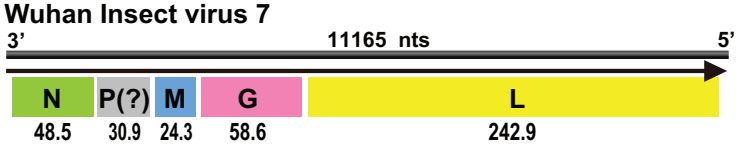


Wuhan Louse Fly Virus 9



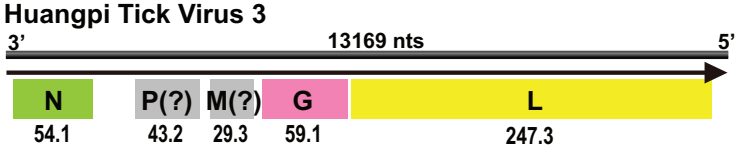
Unclassified Dimarhabdovirus 2

Wuhan Insect virus 7

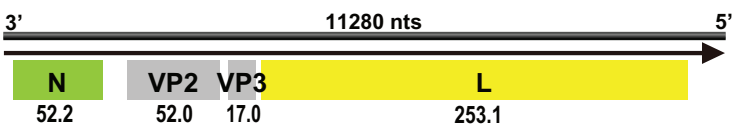


Unclassified Dimarhabdovirus 1

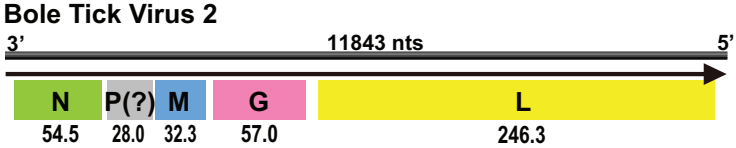
Huangpi Tick Virus 3



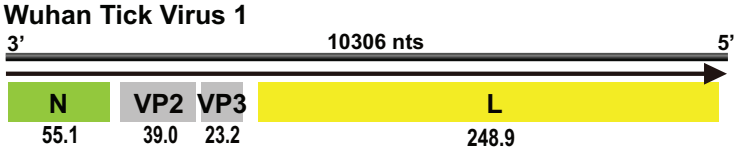
Taishun Tick Virus



Bole Tick Virus 2

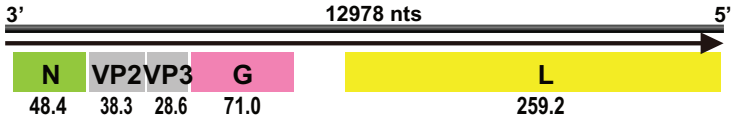


Wuhan Tick Virus 1

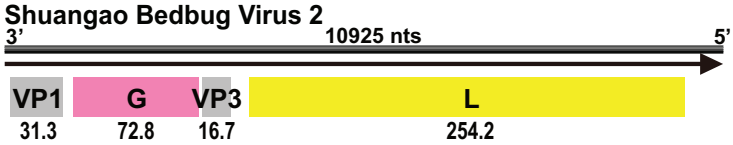


Unclassified Rhabdovirus 2

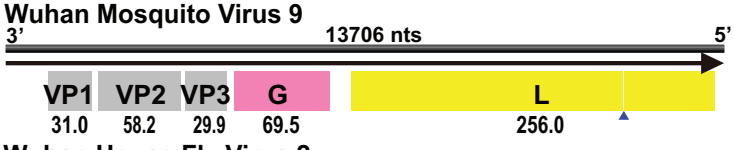
Sanxia Water Strider Virus 5



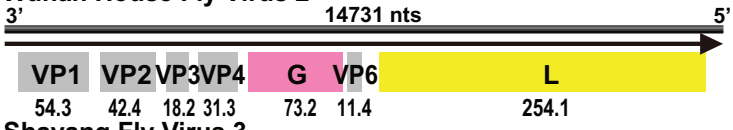
Shuangao Bedbug Virus 2



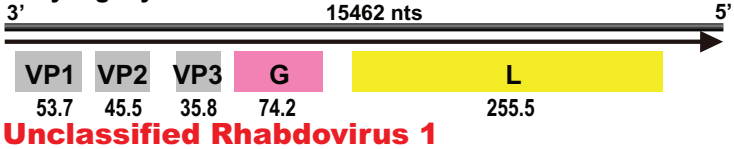
Wuhan Mosquito Virus 9



Wuhan House Fly Virus 2

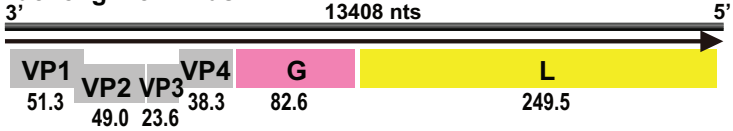


Shayang Fly Virus 3



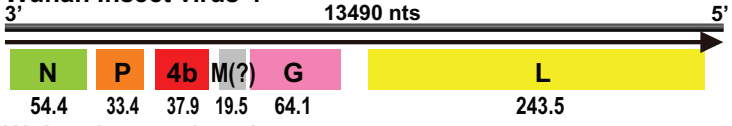
Unclassified Rhabdovirus 1

Tacheng Tick Virus 7

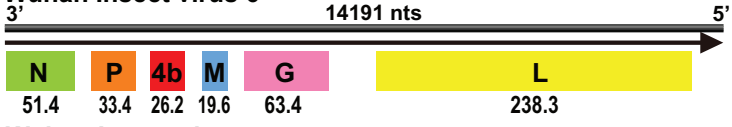


Cytorhabdovirus_like

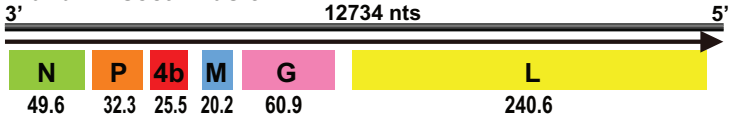
Wuhan Insect virus 4



Wuhan Insect virus 6

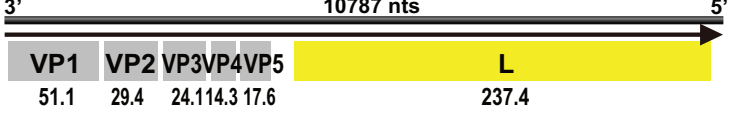


Wuhan Insect virus 5

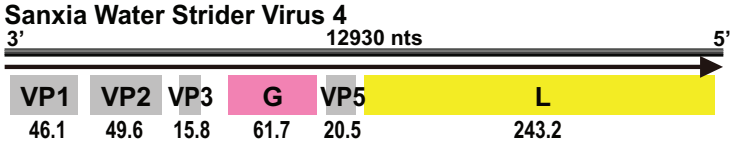


Unclassified Mononegavirus 1

Tacheng Tick Virus 6

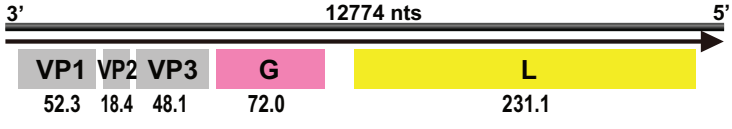


Sanxia Water Strider Virus 4



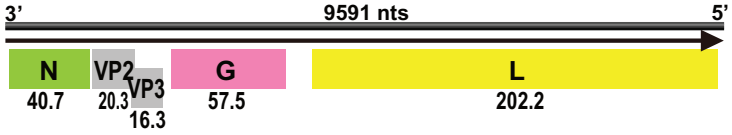
Unclassified Mononegavirus 2

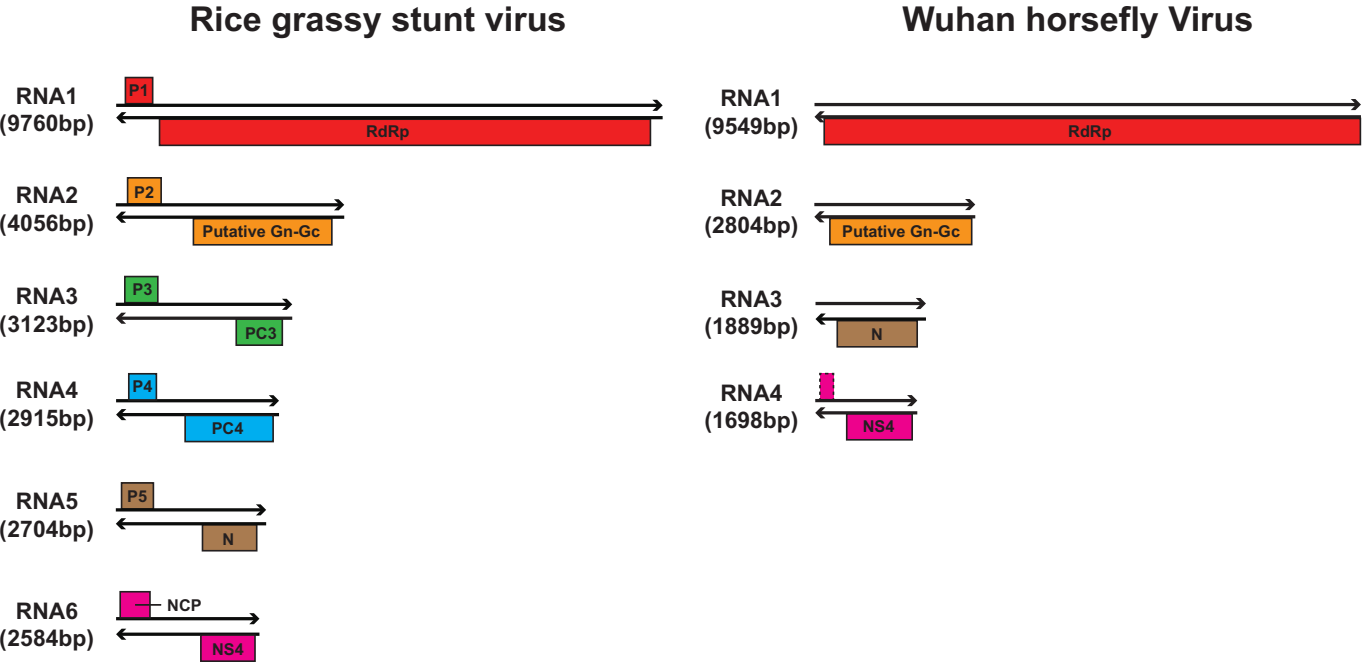
Xincheng Mosquito Virus

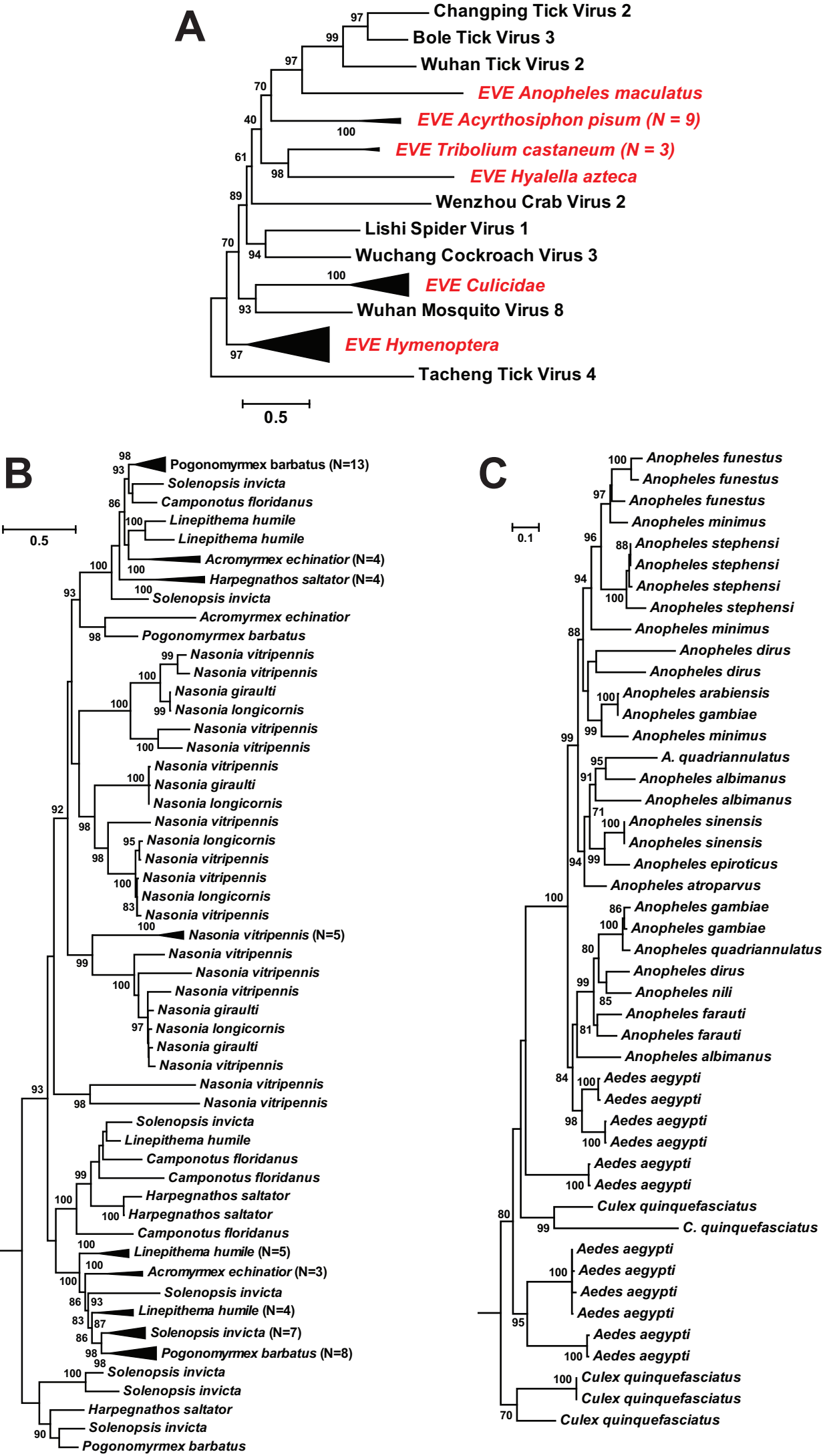


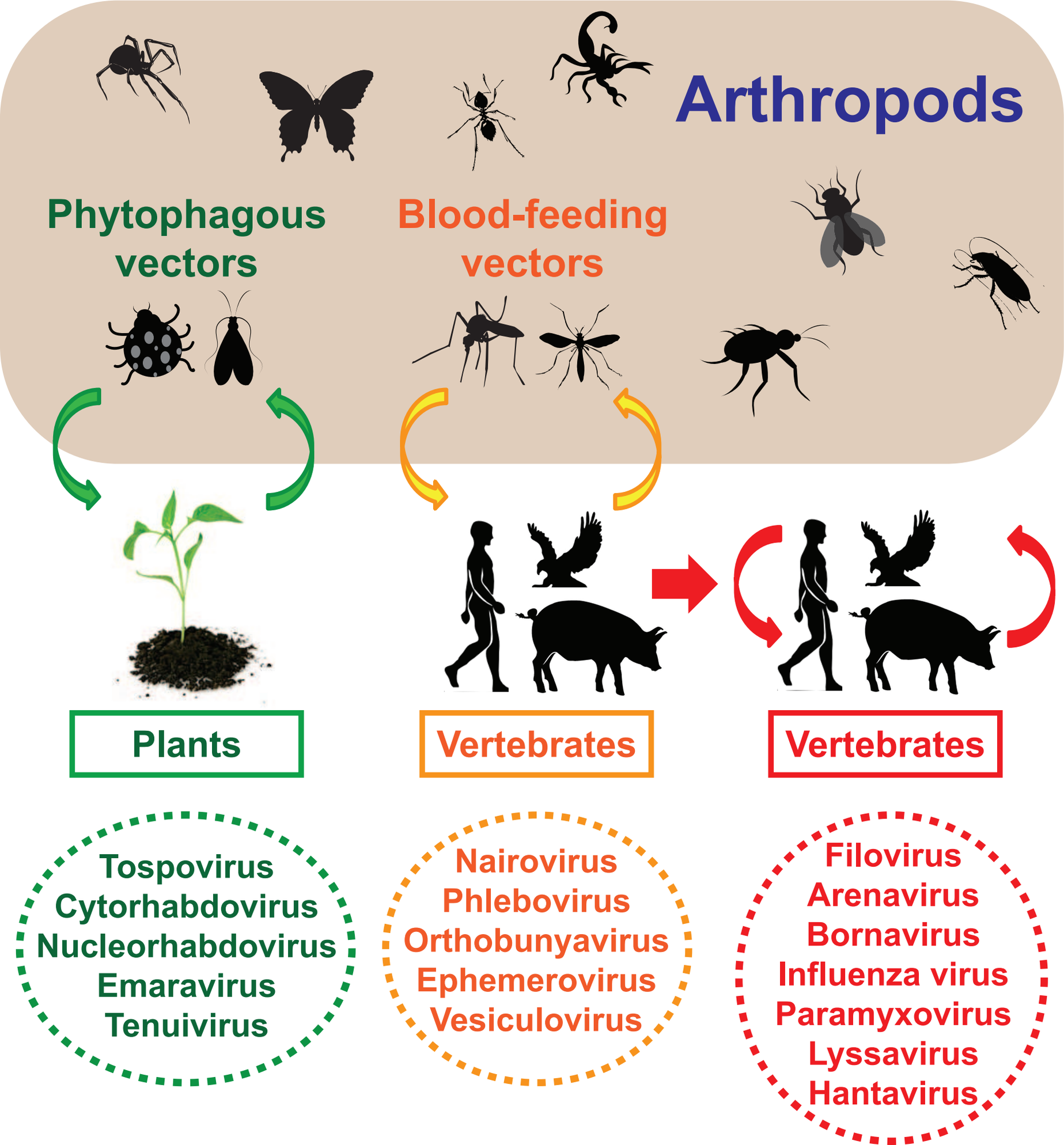
Unclassified Mononegavirus 3

Wenzhou Crab Virus 1




Rice grassy stunt virus
RNA1
(9760bp)
P1
RdRp
RNA2
(4056bp)
P2
Putative Gn-Gc
RNA3
(3123bp)
P3
PC3
RNA4
(2915bp)
P4
PC4
RNA5
(2704bp)
P5
N
RNA6
(2584bp)
NCP
NS4
Wuhan horsefly Virus
RNA1
(9549bp)
RdRp
RNA2
(2804bp)
Putative Gn-Gc
RNA3
(1889bp)
N
RNA4
(1698bp)
NS4


A
Changping Tick Virus 2
Bole Tick Virus 3
Wuhan Tick Virus 2
EVE Anopheles maculatus
EVE Acyrthosiphon pisum (N = 9)
EVE Tribolium castaneum (N = 3)
EVE Hyalella azteca
Wenzhou Crab Virus 2
Lishi Spider Virus 1
Wuchang Cockroach Virus 3
EVE Culicidae
Wuhan Mosquito Virus 8
EVE Hymenoptera
Tacheng Tick Virus 4
0.5
B
0.5
Pogonomyrmex barbatus (N=13)
Solenopsis invicta
Camponotus floridanus
Linepithema humile
Linepithema humile
Acromyrmex echinatior (N=4)
Harpegnathos saltator (N=4)
Solenopsis invicta
Acromyrmex echinatior
Pogonomyrmex barbatus
Nasonia vitripennis
Nasonia vitripennis
Nasonia girault
Nasonia longicornis
Nasonia vitripennis
Nasonia vitripennis
Nasonia vitripennis
Nasonia girault
Nasonia longicornis
Nasonia vitripennis
Nasonia longicornis
Nasonia vitripennis
Nasonia vitripennis
Nasonia longicornis
Nasonia vitripennis
Nasonia vitripennis (N=5)
Nasonia vitripennis
Nasonia vitripennis
Nasonia vitripennis
Nasonia girault
Nasonia longicornis
Nasonia girault
Nasonia vitripennis
Nasonia vitripennis
Nasonia vitripennis
Solenopsis invicta
Linepithema humile
Camponotus floridanus
Camponotus floridanus
Harpegnathos saltator
Harpegnathos saltator
Camponotus floridanus
Linepithema humile (N=5)
Acromyrmex echinatior (N=3)
Solenopsis invicta
Linepithema humile (N=4)
Solenopsis invicta (N=7)
Pogonomyrmex barbatus (N=8)
Solenopsis invicta
Solenopsis invicta
Harpegnathos saltator
Solenopsis invicta
Pogonomyrmex barbatus
C
0.1
Anopheles funestus
Anopheles funestus
Anopheles funestus
Anopheles minimus
Anopheles stephensi
Anopheles stephensi
Anopheles stephensi
Anopheles stephensi
Anopheles minimus
Anopheles dirus
Anopheles dirus
Anopheles arabiensis
Anopheles gambiae
Anopheles minimus
A. quadriannulatus
Anopheles albimanus
Anopheles albimanus
Anopheles sinensis
Anopheles sinensis
Anopheles epiroticus
Anopheles atroparvus
Anopheles gambiae
Anopheles gambiae
Anopheles quadriannulatus
Anopheles dirus
Anopheles nili
Anopheles farauti
Anopheles farauti
Anopheles albimanus
Aedes aegypti
Aedes aegypti
Aedes aegypti
Aedes aegypti
Aedes aegypti
Aedes aegypti
Culex quinquefasciatus
C. quinquefasciatus
Aedes aegypti
Aedes aegypti
Aedes aegypti
Aedes aegypti
Aedes aegypti
Aedes aegypti
Culex quinquefasciatus
Culex quinquefasciatus
Culex quinquefasciatus


Arthropods
Phytophagous vectors
Blood-feeding vectors
Plants
Vertebrates
Vertebrates
Tospovirus
Cytorhabdovirus
Nucleorhabdovirus
Emaravirus
Tenuivirus
Nairovirus
Phlebovirus
Orthobunyavirus
Ephemerovirus
Vesiculovirus
Filovirus
Arenavirus
Bornavirus
Influenza virus
Paramyxovirus
Lyssavirus
Hantavirus