RNA Virus Discovery from *Daphnia* meta-transcriptomes: A novel Tombunoda-like virus based on RNA dependent RNA polymerase identification

Matthew J. Wersebe[1]

[1] School of Biological Sciences, University of Oklahoma, Norman, OK

Correspondence:

mwerseb1@gmail.com

Abstract: (125)

The discovery of RNA viruses from meta-transcriptomes has led to an explosion in viral diversity and identified many novel host-virus associations. To date, no studies have examined the RNA virome of the model zooplankter *Daphnia.* From which only four viruses are known*,* with only one being well understood. Here, I assemble and annotate RNA-dependent RNA polymerase (RdRp) containing contigs from *Daphnia* meta-transcriptomes. Homology searches and phylogenetic inference show that the RdRp genes are related to another crustacean RNA virus identified as order *Tolivirales,* family *Tombusviridae.* A second open reading frame encodes for a protein associated with Nodaviruses also discovered from crustaceans. Taken together, this suggests that Daphnia pulex virus 1 is part of a growing set of tombunodaviruses which are commonly found in invertebrate viromes.

Key Words:

RNA Virus, *Daphnia,* metatranscriptomes, *Tombusviridae*, RNA-dependent RNA Polymerase, RdRp

Introduction:

Recently, the application of meta-transcriptomic ribonucleic acid (RNA) sequencing has led to an explosion in the recognized diversity of RNA viruses from many different hosts and environments (Holmes et al., 2024; Wolf et al., 2020; Shi et al., 2016). Aquatic environments and organisms appear to be particularly biodiverse with respect to the RNA viruses they host, with metagenomic characterization of aquatic habitats routinely yielding many new viruses previously unknown to science (Wolf et al., 2020). Despite the rapid pace at which viruses have been discovered from aquatic ecosystems, little is known about that ecological roles that these pathogens play (e.g., host-parasite interactions) or the impacts that they have on the species commonly found in aquatic communities, such as how much they contributing to mortality or alter ecosystem functioning. This deficit in understanding is particularly striking when considering the paucity of studies describing the "virome", or all the viral elements within freshwater invertebrate species (although see Shi et al., 2016) which occupy key niches in freshwater ecosystems.

Microcrustaceans in the genus *Daphnia* are critical invertebrate species in freshwater zooplankton assemblages. *Daphnia* are considered a keystone species in palustrine food webs, connecting primary production from basal members of the food chain to those at higher trophic levels, such as fish (Lampert and Sommer 2007, Ogorelec et al. 2021). Because of their critical ecological roles in lake and ponds globally, several *Daphnia* species have been well studied (Ebert 2005; 2022), and numerous genomic resources are available for this group (e. g., several chromosome-scale assembles; Wersebe et al. 2023; Barnard-Kubow et al. 2022). *Daphnia* and several fungal (e. g., *Metschnikowia spp.*) and bacterial (e. g., *Pasteuria ramosa*) disease systems are well characterized (Ebert 2005, 2022; Duffy and Sivars-Becker 2005; Decaestecker et al., 2007). Only recently has some attention been turned to understanding potential viral parasites of *Daphnia*. Currently, in the literature there are two characterized viral pathogens of *Daphnia, Daphnia iridescent virus 1* (DIV-1), a large double stranded deoxyribonucleic acid (ds-DNA) virus responsible for White Fat Cell Disease often found in European *Daphnia* (Toenshoff et al., 2018). Additionally, from the United States, *Daphnia mendotae*-associated (Cladocera) hybrid virus (DMClaHV), a small circular single stranded (ss) DNA virus was characterized by metagenomic sequencing from two lakes in New York (Hewson et al., 2013). Hewson and colleagues (2013) showed that there may be mortality associated with infection leading to periodic population declines with potential effects on carbon cycling in lake food webs. Meanwhile, infection with DIV-1 is highly pathogenic and WFCD is often fatal for infected *Daphnia* (Toenshoff et al., 2018).

Evidence of RNA virus infection is incredibly scant from this group. Electron microscopic analysis of some *Daphnia* collected from the Czech Republic show evidence of cytoplasmic polyhedrosis viruses (CPV) a ds-RNA virus (Vavra et al. 2016). Most CPV diversity it known from insects, with 16 recognized species (Matthijnssens et al., 2022). CPV Infections have periodically been reported in the literature, with the first crustacean detection coming from the daphniid *Simocephalus expinosus* (Federici and Hazard, 1975). However, no CPV infections have been genomically characterized to confirm the morphological identifications. Besides these examples, screening of the genome of various *Daphnia* species from the North American *Daphnia pulex* complex has revealed evidence for ancient infections with RNA viruses. Ballinger et al. (2013)

found a set of *Phlebovirus* RNA-dependent RNA Polymerase (RdRp) genes integrated into the genomes of *Daphnia pulex* and other closely related taxa providing the first genomic evidence of paleovirus infections in *Daphnia.* However, a putatively active infection with an RNA virus has remained unconfirmed.

One potential challenge preventing the widespread characterization of viruses from *Daphnia* is a lack of molecular biology resources typically used in virus isolation and research. *Daphnia* do not have immortalized tissue cell cultures systems that are commonly used in viral isolation and research. Attempts have been made to generate primary cell cultures from *Daphnia* (see Robinson et al. 2006), but these showed rapid exhaustion. Renewed efforts have shown promise in generating longer lasting primary cell cultures for ecotoxicological studies (Cp et al. 2025), which may prove useful for virological research. The development of RNA Next Generation Sequencing tools has enabled the rapid characterization of RNA extracted from tissues, enabling the discovery of RNA viruses from a vast array of hosts without the need for virus isolation in cell culture. However, this approach has yet to be applied to *Daphnia.* Here, I investigate the use of RNA meta-transcriptomes from *Daphnia pulex* to identify RNA virus genes. Using bioinformatic approaches, I assemble two RNA virus contigs identified via the annotation of RNA-dependent RNA polymerase (RdRp) genes. Sequence similarity searches and subsequent phylogenetic analysis shows that these sequences belong to the viral order *Tolivirales,* family *Tombusviridae.* The RdRp genes are closely related to another Tombusvirus discovered from a Decapod crustacean from China.

Methods:

I searched the National Center for Biotechnology Information's (NCBI) Sequence Read Archive (SRA) database for RNA-sequencing libraries associated with *Daphnia pulex.* As of October 2025, this search returned of 1,021 unique SRA records. Of this only 8 libraries were identified to belong to samples collected from non-culture conditions. I selected the 8 *Daphnia pulex* RNA sequencing libraires of Hechler et al (2023), indexed under the BioProject PRJNA830892 for bioinformatic discovery of RNA viruses. Hechler et al. (2023), conducted a mesocosm experiment which subjected *Daphnia pulex,* their algal prey, and other natural microbial colonizers to different heating conditions and measured the transcriptomic response to heat stress. Hechler's mock communities included *Daphnia pulex,* and the algal species *Ankistrodesmus falcatus, Scenedesmus quadricauda and Pseudokirchneriella subcapitata,* and any natural colonizers of the experimental tanks. They then subsequently constructed RNA sequencing libraries from *D. pulex* and from the tank water samples. Here, I restrict my analysis to just the RNA sequencing libraries of the *Daphnia* pools.

Bioinformatics Pre-processing:

I downloaded 8 SRA experiments (SRX14976444-SRX14976451) of Hechler and colleagues constructed from their *Daphnia pulex* pools from the European Nucleotide archive SRA FTP server. Next, I quality filtered the FASTQ files using FASTP to remove adapter contamination (using option to automatically detect PE adaptors) and trim low-quality reads (Chen et al., 2018).

Reference Database Construction:

To remove putative host (*Daphnia pulex*) and potentially contaminating reads (from algal community members), I constructed a host and algal reference sequence database for removal by mapping. *Daphnia pulex* has a chromosome-scale reference genome (SAMN22866247) and a mitochondrial genome (Crease, 1999; NC_000844.1). I searched NCBI for reference genome sequences from the other members of the mock communities, I found reference assemblies for *S. quadricauda* (SAMN07360942), *P. subcapitata* (SAMD00078043, aka *Raphidocelis subcapitata*). There was no reference genome for *A. falcatus,* however a there was an unassembled DNA library indexed on SRA reportedly from this species (PRJNA628966; SAMN16512711). Using the approach above, I quality filtered this library using FASTP. I performed *de novo* assembly on the surviving reads using the isolate module of SPADES (Prjibelski et al., 2020). I combined all contigs from these assemblies into a single reference FASTA for read mapping.

Removal of Eukaryotic Reads and Assembly of Unmapped Reads:

Using BBtool's BBmap algorithm (Available: https://sourceforge.net/projects/bbmap), I attempted to map all the RNA reads of Hechler and colleagues against the database consisting of *D. pulex* and the algal species known from the mock communities. I collected all the unmapped RNA seq reads using BBmap's outu option to collect read pairs with no mapping found in the reference database. After mapping was complete, I *de novo* assembled the unmapped reads using the SPADES RNA viral pipeline (Bushmanova et al., 2019).

Identification of Viral RdRp Contigs:

After assembly, I combined all the resulting contigs into a single FASTA file and clustered the files at 98% identity using CD-HIT (Fu et al., 2012) and renamed the clustered sequences by the md5 hash of the retained sequence using a custom R script. Afterwards, I annotated open reading frames in the sequences using the PROKKA annotation pipeline using the kingdom flag set to Viruses (Seemann 2014). After PROKKA annotation, I parsed the output files to identity any potential RNA-dependent RNA-polymerase genes. I then used NCBI's BLASTp tool to search RdRp sequences against the ClusteredNR database to confirm their identification and assign viral taxonomic status (Camancho et al., 2009).

Maximum Likelihood Phylogenetics:

I downloaded the RdRp amino acid multiple sequence alignment (MSA) of the top BLASTp hits in FASTA format using COBALT (Papadopoulos and Agarwala 2007) and renamed the sequence headers using a custom R script. I inferred a maximum likelihood phylogenetic tree in IQTREE (v3) using model finder to identify the best fitting substitution model by Bayesian Information Criteria (BIC) and assessing topological support using 1000 ultra-fast bootstraps (Wong et al., 2025; Hoang et al. 2018). I plotted the tree in R using the package ggtree (Yu 2020).

Results:

Assembly and Annotation:

SPADES Assembly of the unmapped RNA sequencing reads yielded 241,304 contigs ranging in size from 20832 nucleotides in length to 50 nucleotides in length. Clustering at 98% percent identity reduced this to 186,491 contigs which were submitted to PROKKA for annotation. Annotation with PROKKA yielded 45,694 putative transcripts. Of these, two sequences were identified as having homology to other RNA-directed RNA polymerases (aka RNA-dependent) which were both 1335 bp in length. Both contigs contained the RdRp gene had three annotated open reading frames (ORFs) with 2 hypothetical proteins predicted to flank the RdRp sequence. These sequences appear to be identical in the RdRp and one of the other reading frames but are divergent in the third.

BLASTp Sequence Searching:

Using BLASTp to search NCBI's ClusteredNR database revealed that both RdRp sequences (both 444 aa) had a top hit against Qianjiang tombus-like virus 78 (Locus ID: XHA86529.1). While query cover was high (97%), percent identity was low at just 38.5% for both PDGDJCHL_17863 and PDGDJCHL_27476 (evalue: 3e-90; Table 1). Qianjiang tombus-like virus 78 was recovered from meta-transcriptomic profiling of *Procambarus clarkii* (Red Swamp Cray Fish) collected from farmed cray fish in Qianjiang, Hubei Province, China (Guo et al., 2025). BLASTp hits are universally identified as belonging to the riboviruses and often are annotated to either *Tolivirales* or Tombusvirus(-like). Hosts are diverse, including crustaceans (i.e., Red Swamp Cray fish), other arthropods (e.g., *Thrips tabaci*) and plants (e.g., Japanese Iris), representing the diversity of Tombusvirus hosts that have been previously described.

BLASTp searches of the hypothetical proteins from locus tags PDGDJCHL_17862 (459 aa) and PDGDJCHL_27477 (415aa) produced no hits with significant homology when searching against ClusterNR. Sequence searches of PDGDJCHL_17864 and PDGDJCHL_27475 (both 439 aa) revealed congruent results, with homology to both Nodaviruses (again from Red Cray fish) and other some Tombus-like viruses (Table 2). Among the top 100 hits (ranked 15) for these proteins is Qianjiang tombus-like virus 78, with 91.1% query cover and but low percent identity (29.29%). The top hit for both proteins was Qianjiang noda-like virus 94, with a moderate query cover of 73% and a modest (but higher) percent identity of 34.31%. Qianjiang tombus-like virus 78 Metagenome Assembled Genome has a similar recovered structure to the viral MAGs recovered from *Daphnia pulex* libraries, three ORFs with two hypothetical proteins flanking the RdRp (Figure 1A). This genome organization is consistent with other tombunoda-like viruses.

Phylogenetic Inference:

BIC analysis with model finder found that the best fitting model of evolution for the alignment was Q.PFAM+F+I+R6. The recovered RdRp genes are identical to one another in Amino acid composition. The *Daphnia pulex* associated RdRp genes cluster with four other RdRp genes with high bootstrap support (93%; Figure 1B). They are placed ancestral to the branch leading to Qianjiang tombus-like virus 78 (XHA86529.1), the top hit from the ClusterNR BLAST search. The other RdRp within this clade include those discovered from birds (WKV33167.1, WKV33715.1; Lu et al., 2024) and from soil (QDH89441.1; Starr et al., 2019). This clade is the least diverse of all the clades present in the midpoint rooted tree.

Discussion:

The results here, show evidence of a novel RNA virus from a *Daphnia* host. Searching 8 publicly available meta-transcriptomes collected from semi natural conditions led to the discovery of two contigs of putative viral origin by prediction of open reading frames annotated as RNA-dependent RNA polymerases (RdRp). Sequence similarity searches shows that these RdRp sequences belong to the family *Tolivirales,* genus *Tombusviridae.* Phylogenetic analysis shows that these viruses are closely related to diverse RdRp sequences discovered from hosts including a decapod crustacean, birds, and soils. The gene sequences have homology to other RdRps discovered from aquatic habitats (rivers) and many arthropod and invertebrate hosts. Given this initial discovery of "*Daphnia pulex virus 1",* from RNA sequencing data suggests that there may be many yet to be described viruses in natural populations of *Daphnia.* Meta-transcriptomic and meta-genomic surveys of aquatic ecosystems have revealed a huge diversity of both DNA and RNA viruses (Wolf et al. 2020). While the number of studies describing species specific viromes and environment specific virospheres are increasing, very few have been directed at invertebrates. Interestingly, one of the largest RNA virome studies, Shi et al. (2016), included large scale sequencing of many invertebrates throughout China. Reportedly, their freshwater crustacean pool included two daphniid *D. magna* and *D. carinata* however they reported no *Daphnia* specific viruses and only a single Tombus-noda like virus from this pool. Meanwhile, work from Decapod crustaceans shows a huge diversity of Tombus-like viruses and indeed one of the closest related RdRp genes available on NCBI to Daphnia pulex virus 1 was isolated from a decapod crustacean (Guo et al., 2025).

It remains unknown if infection with Daphnia pulex virus 1 has detectable fitness effects for *Daphnia.* Hewson et al (2013) noted that epidemics of DMClaHV may have led to later host abundance declines, presumably from increased mortality or morbidity. Similarly, infection with DIV-1 leading to WFCD is often fatal for infected *Daphnia* and DIV-1 is highly virulent as noted by Toenshoff et al. (2018) citing Green (1957). While not a guarantee, host-parasite interactions may evolve to more a benign steady state if they co-exist over many generations (May and Anderson, 1990). The *Daphnia* used in this investigation were collected from a wild population and successfully established a population in the experimental units and further survived prolonged exposure to heat stress (Hechler et al. 2023). This would suggest that infection with Daphnia pulex virus 1 may not represent an acute or even fatal infection, rather an infection that *Daphnia* can persist with and survive even during stressful periods. Additionally, other tombus-like viruses closely related to the one I identified were discovered from, farmed crayfish, a Decapod crustacean (Guo et al., 2025). One interpretation of both these lines of evidence is that Daphnia pulex virus 1 has a long history with *Daphnia* having evolved a steady state where infection is largely benign. Additional sampling, basic natural history studies, and infection assays in the laboratory would need to be conducted to better ascertain if this hypothesis is indeed supported. Additionally, this would also help to establish whether we can expect *Daphnia* to clear infection or if infection remains chronic as with some other RNA viral pathogens such as Hepatitis C virus (HCV; Thomas et al., 2000) or Human Immunodeficiency Virus (HIV; Lyes et al., 2000), only to cause increased morbidity and mortality after a long period. Furthermore, it would be key to know if RNA viruses are horizontally transmitted across clonal or sexual generations in *Daphnia* and how to this relates to clonal fitness and diapause success. However, whether Daphnia pulex virus 1 is in fact infecting the *D. pulex* used in this study is also key to understand and should be further explored.

There exists a possibility that this viral element is part of the larger holobiome of the animals or is possibly associated with the algal food within the gut of the experimental animals and not infecting the *Daphnia* at all. This, however, seems unlikely as the Tombusviridae is not known to infect green algae or other microbes, rather it is associated almost exclusively with terrestrial plants and arthropods (White and Nagy 2004; Lozier et al., 2024). Association with *Daphnia* is supported by the BLASTp and follow up phylogenetic analysis which shows close allyship with angiosperm, chordate, mollusc and other arthropod viruses. I found no algal hosts among the top blast hits.

The BLASTp hits for the RdRp and the other hypothetical protein (PDGDJCHL_17864 & PDGDJCHL_27475) had differing taxonomies. The RdRp sequence had homology to Tombus-like viruses meanwhile the second protein had homologies to noda-like viruses. This result is not surprising, as Nodaviruses and Tombusvirus are closely allied and appear to horizontally exchange sequences (Shi et al. 2016; Dolja and Koonin 2017) through recombination (Kovalev et al. 2019). Tombunodaviruses were first described by Greninger and DeRisi (2015) from an environmental sample collected in California, USA. Daphnia pulex virus 1 and Tombunodavirus UC1 share a similar RdRp and genome organization. Given these similarities, and the widespread identification of similar Tombunodaviruses from other invertebrate viromes (e.g., Shi et al. 2016), it appears Daphnia pulex virus 1 belongs to this group of newly recognized viruses

Here, I describe an RNA virus, Daphnia pulex virus 1, discovered by mining metatranscriptomes of *Daphnia pulex* and assembling RNA-dependent RNA polymerases *de novo.* Sequence similarity searches and phylogenetic inference reveal the recovered virus is related to viruses infecting angiosperms and recovered from RNA virus assemblages inferred from other arthropods, including crustaceans, molluscs, and chordates. It remains unknown the extent to which this virus is pathogenic to *Daphnia,* and further studies should attempt to further elucidate the virome of *Daphnia* and the extent to which RNA virus infection may cause disease, increased morbidity, or mortality. Given *Daphnia's* central role in aquatic food webs, it remains important to further explore their associated RNA virome to understand if infections influence ecosystem functioning.

Open Research and Data Availability:

I analyzed the RNA sequencing libraries (SRX14976444-SRX14976451) accessioned under BioProject (PRJNA830892). Viral genome sequences have been submitted to NCBI Genbank and are available under accessions PX577120 and PX577121. Scripts and data are provided on GitHub at: https://github.com/mwersebe/DaphniaRNAvirus.

Conflicts of Interest and Funding:

I declare no conflicts of interest. No funding was used to support this research.

Works Cited:

Ballinger MJ, Bruenn JA, Kotov AA, Taylor DJ. Selectively maintained paleoviruses in Holarctic water fleas reveal an ancient origin for phleboviruses. Virology. 2013 Nov;446(1-2):276-82. doi: 10.1016/j.virol.2013.07.032. Epub 2013 Sep 6. PMID: 24074591.

Barnard-Kubow KB, Becker D, Murray CS, Porter R, Gutierrez G, Erickson P, Nunez JCB, Voss E, Suryamohan K, Ratan A, Beckerman A, Bergland AO. Genetic Variation in Reproductive Investment Across an Ephemerality Gradient in Daphnia pulex. Mol Biol Evol. 2022 Jun 2;39(6):msac121. doi: 10.1093/molbev/msac121. PMID: 35642301; PMCID: PMC9198359.

Bushmanova E, Antipov D, Lapidus A, Prjibelski AD. rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. Gigascience. 2019 Sep 1;8(9):giz100. doi: 10.1093/gigascience/giz100. PMID: 31494669; PMCID: PMC6736328.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics. 2009 Dec 15;10:421. doi: 10.1186/1471-2105-10-421. PMID: 20003500; PMCID: PMC2803857.

Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018 Sep 1;34(17):i884-i890. doi: 10.1093/bioinformatics/bty560. PMID: 30423086; PMCID: PMC6129281.

Cp S, Tm MK, Balakrishnan S, Kunjiraman S, Sarasan M, Magnuson JT, Puthumana J. Establishment of a cell culture from Daphnia magna as an in vitro model for (eco)toxicology assays: Case study using Bisphenol A as a representative cytotoxic and endocrine disrupting chemical. Aquat Toxicol. 2025 Jan;278:107173. doi: 10.1016/j.aquatox.2024.107173. Epub 2024 Nov 23. PMID: 39612670.

Crease TJ. The complete sequence of the mitochondrial genome of Daphnia pulex (Cladocera: Crustacea). Gene. 1999 Jun 11;233(1-2):89-99. doi: 10.1016/s0378-1119(99)00151-1. PMID: 10375625.

Decaestecker E, Gaba S, Raeymaekers JA, Stoks R, Van Kerckhoven L, Ebert D, De Meester L. Host-parasite 'Red Queen' dynamics archived in pond sediment. Nature. 2007 Dec 6;450(7171):870-3. doi: 10.1038/nature06291. Epub 2007 Nov 14. PMID: 18004303.

Dolja VV, Koonin EV. Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. Virus Res. 2018 Jan 15;244:36-52. doi: 10.1016/j.virusres.2017.10.020. Epub 2017 Nov 8. PMID: 29103997; PMCID: PMC5801114.

Duffy MA, Sivars-Becker L. Rapid evolution and ecological host-parasite dynamics. Ecol Lett. 2007 Jan;10(1):44-53. doi: 10.1111/j.1461-0248.2006.00995.x. PMID: 17204116.

Ebert D. Ecology, Epidemiology, and Evolution of Parasitism in Daphnia [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2005. Available from: https://www.ncbi.nlm.nih.gov/books/NBK2036/

Ebert D. Daphnia as a versatile model system in ecology and evolution. Evodevo. 2022 Aug 8;13(1):16. doi: 10.1186/s13227-022-00199-0. PMID: 35941607; PMCID: PMC9360664.

Federici, B., Hazard, E. Iridovirus and cytoplasmic polyhedrosis virus in the freshwater daphnid Simocephalus expinosus. Nature 254, 327–328 (1975). https://doi.org/10.1038/254327a0

Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012 Dec 1;28(23):3150-2. doi: 10.1093/bioinformatics/bts565. Epub 2012 Oct 11. PMID: 23060610; PMCID: PMC3516142.

Guo G, Liu Z, Zeng J, Yan H, Chen G, Han P, He X, Zhou D, Weng S, He J, Wang M. Virome analysis unveils a rich array of newly identified viruses in the red swamp crayfish Procambarus clarkii.

Virology. 2025 Jan;601:110308. doi: 10.1016/j.virol.2024.110308. Epub 2024 Nov 16. PMID: 39556981.

Green J, 1974 Parasites and epibionts of Cladocera. Trans. Zool. Soc. Lond. 32(6): 417–515. 10.1111/j.1096-3642.1974.tb00031.x

Greninger AL, DeRisi JL. Draft Genome Sequence of Tombunodavirus UC1. Genome Announc. 2015 Jul 2;3(4):e00655-15. doi: 10.1128/genomeA.00655-15. PMID: 26139709; PMCID: PMC4490837.

Hechler RM, Yates MC, Chain FJJ, Cristescu ME. Environmental transcriptomics under heat stress: Can environmental RNA reveal changes in gene expression of aquatic organisms? Mol Ecol. 2025 Jul;34(13):e17152. doi: 10.1111/mec.17152. Epub 2023 Oct 4. PMID: 37792902; PMCID: PMC12186723.

Hewson Ian, Ng Gabriel, Li WenFang, LaBarre Brenna A., Aguirre Isabel, Barbosa Jorge G., Breitbart Mya, Greco Anthony W., Kearns Colleen M., Looi Alexander, Schaffner Lindsay R., Thompson Philip D., Hairston Jr. Nelson G., (2013), Metagenomic identification, seasonal dynamics, and potential transmission mechanisms of a Daphnia-associated single-stranded DNA virus in two temperate lakes, Limnology and Oceanography, 58, doi: 10.4319/lo.2013.58.5.1605.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol. 2018 Feb 1;35(2):518-522. doi: 10.1093/molbev/msx281. PMID: 29077904; PMCID: PMC5850222.

Holmes EC, Krammer F, Goodrum FD. Virology-The next fifty years. Cell. 2024 Sep 19;187(19):5128-5145. doi: 10.1016/j.cell.2024.07.025. PMID: 39303682; PMCID: PMC11467463.

Kovalev N, Pogany J, Nagy PD. Interviral Recombination between Plant, Insect, and Fungal RNA Viruses: Role of the Intracellular Ca2+/Mn2+ Pump. J Virol. 2019 Dec 12;94(1):e01015-19. doi: 10.1128/JVI.01015-19. PMID: 31597780; PMCID: PMC6912095.

Lampert W., U. Sommer, Limnoecology: The Ecology of Lakes and Streams (Oxford University Press, 2007)

Lozier, Z., Hill, L., Semmann, E., & Miller, W. A. (2024). A proposed new Tombusviridae genus featuring extremely long 5’ untranslated regions and a luteo/polerovirus-like gene block. Frontiers in Virology, 4, 1422934. https://doi.org/10.3389/fviro.2024.1422934

Lyles RH, Muñoz A, Yamashita TE, Bazmi H, Detels R, Rinaldo CR, Margolick JB, Phair JP, Mellors JW. Natural history of human immunodeficiency virus type 1 viremia after seroconversion and proximal to AIDS in a large cohort of homosexual men. Multicenter AIDS Cohort Study. J Infect Dis. 2000 Mar;181(3):872-80. doi: 10.1086/315339. PMID: 10720507.

May RM, Anderson RM. Parasite-host coevolution. Parasitology. 1990;100 Suppl:S89-101. doi: 10.1017/s0031182000073042. PMID: 2122393.

Ogorelec, Z. , C. Wunsch, A. J. Kunzmann, P. Octorina, J. I. Navarro, Large daphniids are keystone species that link fish predation and phytoplankton in trophic cascades. Fundam. Appl. Limnol. 194, 297–309 (2021).

Papadopoulos JS, Agarwala R. COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics. 2007 May 1;23(9):1073-9. doi: 10.1093/bioinformatics/btm076. Epub 2007 Mar 1. PMID: 17332019.

Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. Curr Protoc Bioinformatics. 2020 Jun;70(1):e102. doi: 10.1002/cpbi.102. PMID: 32559359.

Robinson CD, Lourido S, Whelan SP, Dudycha JL, Lynch M, Isern S. Viral transgenesis of embryonic cell cultures from the freshwater microcrustacean Daphnia. J Exp Zool A Comp Exp Biol. 2006 Jan 1;305(1):62-7. doi: 10.1002/jez.a.250. PMID: 16358270; PMCID: PMC4863945.

Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ. Redefining the invertebrate RNA virosphere. Nature. 2016 Dec 22;540(7634):539-543. doi: 10.1038/nature20167. Epub 2016 Nov 23. PMID: 27880757.

Starr EP, Nuccio EE, Pett-Ridge J, Banfield JF, Firestone MK. Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. Proc Natl Acad Sci U S A. 2019 Dec 17;116(51):25900-25908. doi: 10.1073/pnas.1908291116. Epub 2019 Nov 26. PMID: 31772013; PMCID: PMC6926006.

Thomas DL, Astemborski J, Rai RM, Anania FA, Schaeffer M, Galai N, Nolt K, Nelson KE, Strathdee SA, Johnson L, Laeyendecker O, Boitnott J, Wilson LE, Vlahov D. The natural history of hepatitis C virus infection: host, viral, and environmental factors. JAMA. 2000 Jul 26;284(4):450-6. doi: 10.1001/jama.284.4.450. PMID: 10904508.

Toenshoff ER, Fields PD, Bourgeois YX, Ebert D. The End of a 60-year Riddle: Identification and Genomic Characterization of an Iridovirus, the Causative Agent of White Fat Cell Disease in Zooplankton. G3 (Bethesda). 2018 Mar 28;8(4):1259-1272. doi: 10.1534/g3.117.300429. PMID: 29487186; PMCID: PMC5873915.

Vávra J, Bílý T, Nebesářová J, Federici BA. Occurrence, pathology, and ultrastructure of iridovirus and cytoplasmic polyhedrosis viruses in daphnids from the Czech Republic. J Invertebr Pathol. 2016 Oct;140:35-38. doi: 10.1016/j.jip.2016.07.008. Epub 2016 Jul 19. PMID: 27449679.

Wersebe MJ, Sherman RE, Jeyasingh PD, Weider LJ. The roles of recombination and selection in shaping genomic divergence in an incipient ecological species complex. Mol Ecol. 2023 Mar;32(6):1478-1496. doi: 10.1111/mec.16383. Epub 2022 Feb 15. PMID: 35119153.

White KA, Nagy PD. Advances in the molecular biology of tombusviruses: gene expression, genome replication, and recombination. Prog Nucleic Acid Res Mol Biol. 2004;78:187-226. doi: 10.1016/S0079-6603(04)78005-8. PMID: 15210331.

Wolf YI, Silas S, Wang Y, Wu S, Bocek M, Kazlauskas D, Krupovic M, Fire A, Dolja VV, Koonin EV. Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. Nat Microbiol. 2020 Oct;5(10):1262-1270. doi: 10.1038/s41564-020-0755-4. Epub 2020 Jul 20. PMID: 32690954; PMCID: PMC7508674.

T.K.F. Wong, N. Ly-Trong, H. Ren, H. Banos, A.J. Roger, E. Susko, C. Bielow, N. De Maio, N. Goldman, M.W. Hahn, G. Huttley, R. Lanfear, B.Q. Minh (2025) IQ-TREE 3: Phylogenomic Inference Software using Complex Evolutionary Models. Submitted, https://doi.org/10.32942/X2P62N.

Yu G. Using ggtree to Visualize Data on Tree-Like Structures. Curr Protoc Bioinformatics. 2020 Mar;69(1):e96. doi: 10.1002/cpbi.96. PMID: 32162851.

Figures:

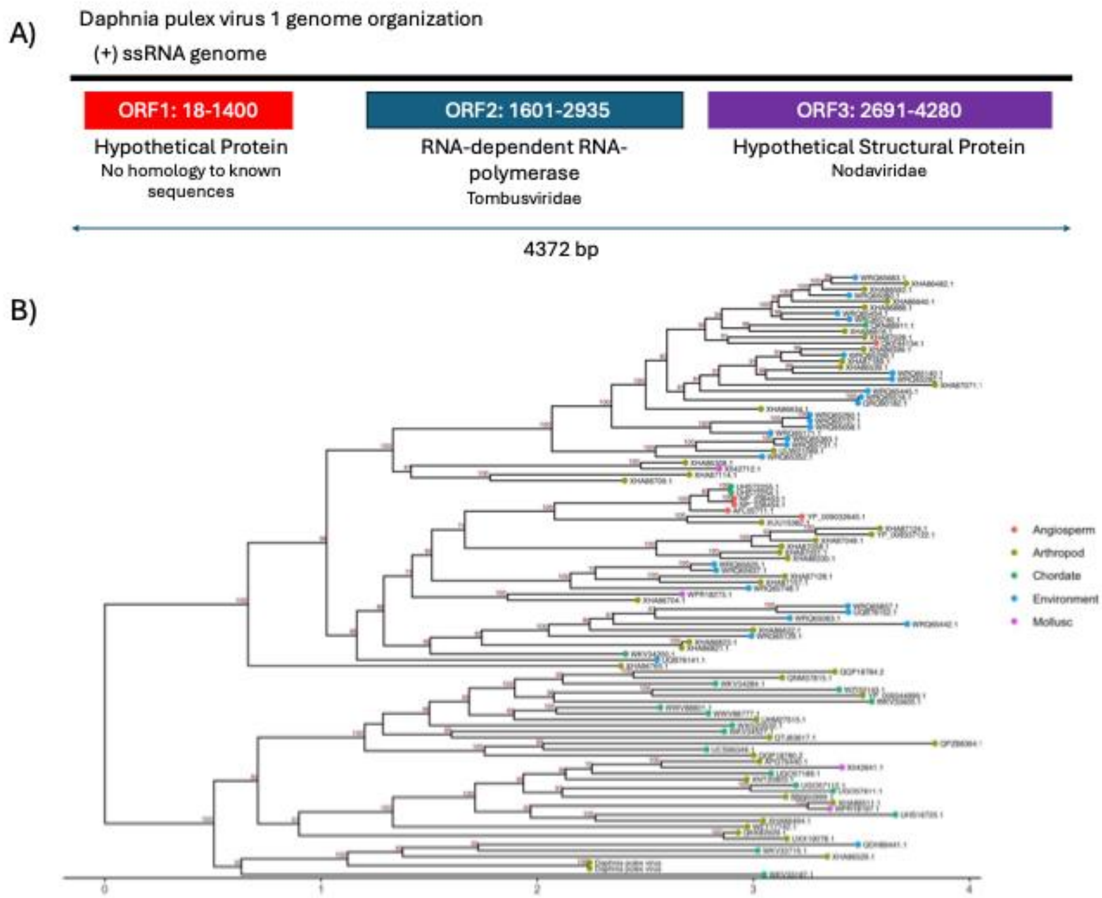

Figure 1: **A)** Genomic organization of Daphnia pulex virus 1. The positive sense single stranded RNA genome [(+)ss RNA], has three open reading frames (ORF) when translated with the standard genetic code. ORF1 has no significant homology to known sequences. ORF2 is the RdRp gene and is closely related to Tombusviridae RdRps. ORF 3 has homology to Nodaviridae structural proteins. **B)** Midpoint rooted maximum likelihood phylogeny of Daphnia pulex virus 1 RNA-dependent RNA polymerase amino acid sequences and the top 100 BLASTp hits. Tip are colored according to the host phylum/clade or labelled as environment if assembled from an environmental sample. BLASTp hits were primarily from arthropods (Tan) or from environmental sources (Blue). Environmental samples primarily included riverine and aquatic habitats. In addition to arthropods, hosts of other closely allied viruses included chordates (Green; birds, mammals and fish) and Molluscs (Pink; freshwater mussels). Node labels indicate bootstrap support estimated with 1000 ultra-fast bootstraps.

Tables:

ACCESSION	CLUSTER NAME	SIZE	EVALUE	PID (%)	QUERY COVERY (%)	SOURCE	SOURCE LOCATION
XHA86529.1	Qianjiang tombus-like virus 78	1	3e-90	38.58	97	Red Crayfish	China
WKV33715.1	Riboviria sp.	1	6e-68	38.10	94	Warbler	China
QDH89441.1	Riboviria sp.	1	9e-70	41.27	75	Soil	USA
XII42641.1	clirnapec virus 39	1	7e-70	36.25	87	Freshwater mussel	USA
XHA86494.1	Qiangjiang tombus-like virus 66	1	6e-67	35.61	85	Red Crayfish	China
XUU15382.1	Carmovirus sp.	1	4e-65	38.24	83	Honeybee	Uzbekistan
APG76440.1	Sanxia water strider virus 14	1	5e-65	35.46	87	Freshwater Atyid shrimp	China
QTJ63617.1	Phasmatodean tombus-related virus	1	8e-64	37.30	82	Leaf insect	Germany (insect culture)
UXX19078.1	Fangzheng tombus-like virus	1	1e-62	37.31	85	Taiga tick	China
XHA86821.1	Qianjiang tombus-like virus 183	1	3e-62	34.78	84	Red Crayfish	China

Table 1: Top 10 BLASTp hit of the RdRp sequences from ClusteredNR database. Sequences are shorted by BLASTp evalue.

ACCESSION	CLUSTER NAME	SIZE	EVALUE	PID (%)	QUERY COVERY (%)	SOURCE	SOURCE LOCATION
XHA86089.1	Qianjiang noda-like virus 94	1	9e-46	34.31	73	Red Crayfish	China
WPR18221.1	Crogonang virus 45	1	4e-33	27.72	87	Freshwater mussel	USA
WZH61881.1	Sopobep virus	1	7e-33	28.46	79	Farm soil	Australia
APG76166.1	Beihai noda-like virus 6	1	2e-33	30.72	73	Penaeid shrimp	China
XHA85966.1	Qianjiang noda-like virus 37	1	9e-31	29.92	79	Red Crayfish	China
XKB76411.1	Nodamuvirales sp.	1	8e-27	37.56	73	River water	Germany
WRQ64796.1	Sobelivirales sp.	1	9e-26	26.03	47	River water	Germany
XKB76231.1	Nodamuvirles sp.	1	2e-24	26.03	97	River water	Germany
WRQ65002.1	Tolivirales sp.	1	2e-23	27.50	98	River water	Germany
QDH89441.1	Riboviria sp.	1	2e-23	29.41	75	Soil	USA

Table 2: Top 10 BLASTp hits for the hypothetical structural proteins from ORF3. Sequences are shorted by BLASTp evalue.