

Metagenomic sequencing reveals a lack of virus exchange between native and invasive freshwater fish across the Murray–Darling Basin, Australia

Vincenzo A. Costa,¹ Jonathon C.O. Mifsud,^{1,†} Dean Gilligan,²
Jane E. Williamson,¹ Edward C. Holmes,^{3,‡} and Jemma L. Geoghegan^{1,4,5,*,&}

¹Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia, ²NSW Department of Primary Industries, Batemans Bay Fisheries Office, Batemans Bay 2536, Australia, ³Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Life and Environmental Sciences and Sydney Medical School, The University of Sydney, Sydney, NSW 2006, Australia, ⁴Department of Microbiology and Immunology, University of Otago, Dunedin 9016, New Zealand and ⁵Institute of Environmental Science and Research, Wellington, Porirua 5022, New Zealand

*Corresponding author: E-mail: jemma.geoghegan@otago.ac.nz

†<https://orcid.org/0000-0001-8934-9193>

‡<https://orcid.org/0000-0001-9596-3552>

§<https://orcid.org/0000-0003-0970-0153>

Abstract

Biological invasions are among the biggest threats to freshwater biodiversity. This is increasingly relevant in the Murray–Darling Basin, Australia, particularly since the introduction of the common carp (*Cyprinus carpio*). This invasive species now occupies up to ninety per cent of fish biomass, with hugely detrimental impacts on native fauna and flora. To address the ongoing impacts of carp, cyprinid herpesvirus 3 (CyHV-3) has been proposed as a potentially effective biological control agent. Crucially, however, it is unknown whether CyHV-3 and other cyprinid herpesviruses already exist in the Murray–Darling. Further, little is known about those viruses that naturally occur in wild freshwater fauna, and the frequency with which these viruses jump species boundaries. To document the evolution and diversity of freshwater fish viromes and better understand the ecological context to the proposed introduction of CyHV-3, we performed a meta-transcriptomic viral survey of invasive and native fish across the Murray–Darling Basin, covering over 2,200 km of the river system. Across a total of thirty-six RNA libraries representing ten species, we failed to detect CyHV-3 nor any closely related viruses. Rather, meta-transcriptomic analysis identified eighteen vertebrate-associated viruses that could be assigned to the *Arenaviridae*, *Astroviridae*, *Bornaviridae*, *Caliciviridae*, *Coronaviridae*, *Chuviridae*, *Flaviviridae*, *Hantaviridae*, *Hepeviridae*, *Paramyxoviridae*, *Picornaviridae*, *Poxviridae*, *Reoviridae* and *Rhabdoviridae* families, and a further twenty-seven that were deemed to be associated with non-vertebrate hosts. Notably, we revealed a marked lack of viruses that are shared among invasive and native fish sampled here, suggesting that there is little virus transmission from common carp to native fish species, despite co-existing for over

fifty years. Overall, this study provides the first data on the viruses naturally circulating in a major river system and supports the notion that fish harbour a large diversity of viruses with often deep evolutionary histories.

Key words: virome; meta-transcriptomics; fish virus; freshwater fish; virus evolution; RNA sequencing.

1. Introduction

Anthropogenic stressors such as pollution, climate change and the introduction of exotic species continue to pose a significant threat to freshwater habitats, with almost one-third of all fish species threatened by extinction (World Wildlife Fund 2021). The Murray–Darling Basin, the largest freshwater river system in Australia, harbours at least twelve exotic freshwater fish species (Barrett, Bamford, and Jackson 2014). Key among these are eastern mosquitofish (*Gambusia holbrooki*), redfin perch (*Perca fluviatilis*) and, most notably, common carp (*Cyprinus carpio*). Common carp (also known as European carp) were initially introduced into Australia during the mid-1800s for aquaculture operations and again on several occasions throughout the 1900s (Koehn 2004; Forsyth et al. 2013). During extensive flooding events during the 1970s, carp spread across much of the basin and now represent up to ninety per cent of total fish biomass (Forsyth et al. 2013).

The invasion of carp has been hugely detrimental to Australian freshwater ecosystems (Koehn 2004; Forsyth et al. 2013). Impacts include increased water turbidity, decreased light penetration, erosion of riverbanks, changes in the abundance and diversity of native invertebrate communities and outcompeting native fish species for habitat and resources (Koehn 2004; Forsyth et al. 2013; Vilizzi et al. 2014). Several control methods have been proposed to control invasive carp; nevertheless, their resilience and high fecundity create significant challenges (Hayes et al. 2014). This has stimulated research into biological control methods, such as deployment of the virus cyprinid herpesvirus 3 (CyHV-3) (McColl et al. 2017; McColl, Cooke, and Sunarto 2014).

CyHV-3 is a double-stranded DNA virus (family *Alloherpesviridae*, order *Herpesvirales*) first isolated from farmed carp in the late 1990s (Matsui et al. 2008). Since its discovery, it has been responsible for large disease outbreaks worldwide with a mortality rate of up to 80 per cent in domestic carp (Michel et al. 2010). CyHV-3 is transmitted horizontally through direct contact with skin lesions or secretion of viral particles in freshwater where it can survive for up to 3 days (Shimizu et al. 2006). The host range of CyHV-3 is currently limited to koi and common carp (Michel et al. 2010). While CyHV-3 DNA has been identified in goldfish (*Carrasius auratus*) (Bergmann et al. 2010), it is still relatively unclear whether infection occurs in these species (Ilouze et al. 2011; Tolo et al. 2021).

Initial laboratory trials suggest that CyHV-3 is safe for non-target species (McColl et al. 2017). However, little is known about the viruses that naturally circulate in Australian native freshwater fauna, including any prior evidence for the existence of CyHV-3 (Kopf et al. 2019), nor on the time-scales and frequency with which viruses jump between fish hosts. To completely assess the safety and efficacy of any virus biocontrol agent, including CyHV-3, a comprehensive assessment of the viruses that naturally infect both native and invasive species is required.

Following the advent of meta-transcriptomic sequencing, it is now possible to characterize the entire set of viruses—the virome—within a given host (Shi, Zhang, and Holmes 2018;

Zhang, Shi, and Holmes 2018). Fish, in particular, harbour a high abundance and diversity of viruses often with deep evolutionary histories (Shi et al. 2018; Zhang et al. 2018). However, despite the antiquity and diversity of fish viruses, there are few studies of virus diversity and evolution in wild freshwater fish populations, particularly in the context of biological invasions.

Determining the viromes of invasive freshwater fish like the common carp will enhance our understanding of the broad-scale factors that influence virus emergence and evolution. As the date and site of their introduction is well-documented in Australian waters, these species can potentially provide important information on the both rate of cross-species transmission and how frequently viruses might move between invasive and native species. In addition, despite representing a small fraction of the earth's surface water, freshwater environments serve as a habitat for forty to fifty per cent of total fish species, harbouring the greatest biodiversity per land area (Dudgeon et al. 2006). Such habitats are subject to rapid environmental change, which may significantly impact species connectivity (Johnson and Paull 2011). Since contact and exposure between hosts are vital for cross-species transmission of viruses (Parrish et al. 2008), these species may also inform us on the ecological factors that impact virome composition within a given host.

We performed a meta-transcriptomic viral survey of invasive and native freshwater fish species across the Murray–Darling Basin in Australia to document the diversity and evolution of freshwater fish viromes and, from this, better understand the ecological drivers of virus evolution and emergence. To the best of our knowledge, this is the largest survey of freshwater fish viruses undertaken to date. In particular, we aimed to determine whether CyHV-3 is already present in common carp in Australia (Kopf et al. 2019), and whether there is evidence for transmission of existing viruses between exotic and native species. As such, we provide important information on the ecological and evolutionary context for the potential release of future virus biocontrols.

2. Methods

2.1 Ethics

Fish sampling was conducted with animal ethics approval (ref: 2019/035) from the Animal Ethics Committee (AEC) at Macquarie University, Sydney, NSW. Biosafety was approved by Macquarie University (ref: 5201700856).

2.2 Sample Collection

We compared the viromes of native and invasive fish species occupying different areas across the Murray–Darling Basin, Australia (Fig. 1). Sampling occurred between January and March 2020. A total of seven native fish species were collected: bony herring (*Nematalosa erebi*) (n = 20), spangled perch (*Leiopotherapon unicolor*) (n = 1), Australian smelt (*Retropinna semoni*) (n = 12), Murray–Darling rainbowfish (*Melanotaenia fluviatilis*) (n = 17), flat-headed gudgeon (*Philypnodon grandiceps*) (n = 9), western carp-gudgeon (*Hypseleotris* spp.) (n = 20) and unspotted

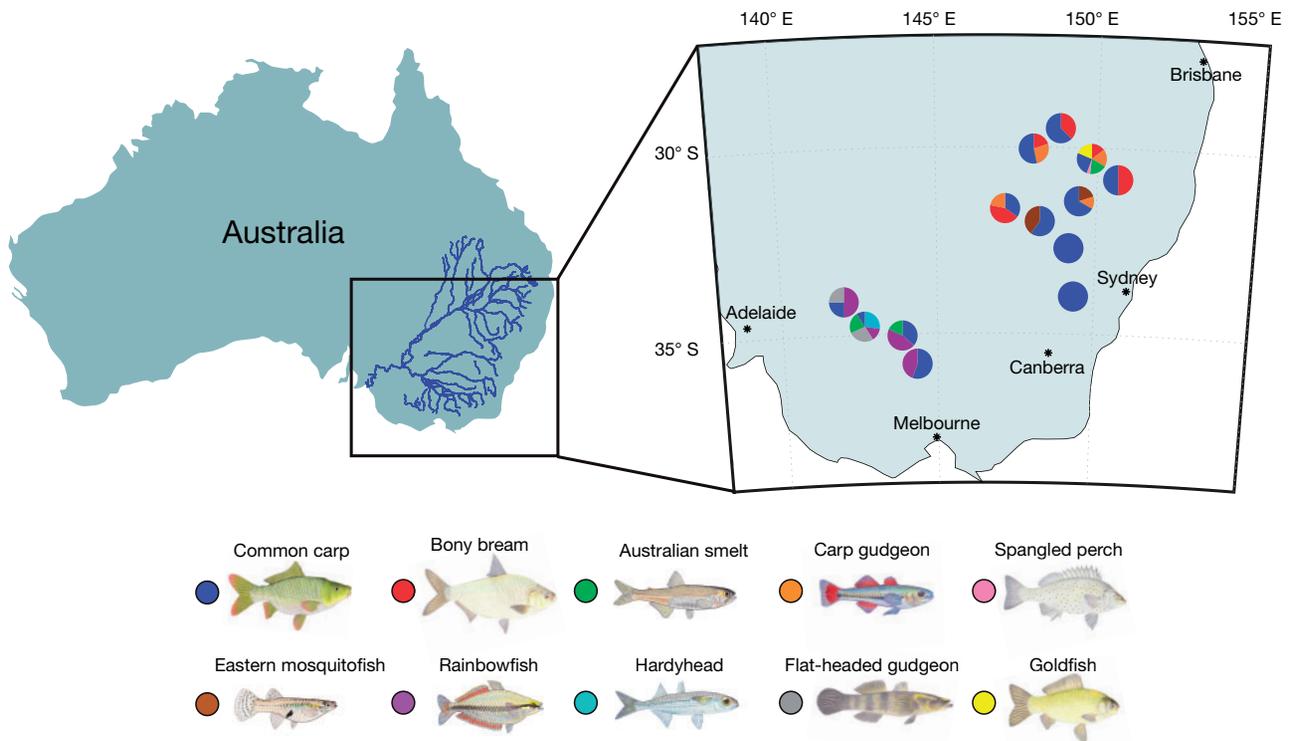


Figure 1. Fish sampling locations. Map illustrating freshwater sampling locations across the Murray–Darling Basin, Australia. Pie charts show the abundance and diversity of fish species captured at each site. Colours correspond to the fish species sampled as illustrated below the map.

hardyhead (*Craterocephalus fulvus*) ($n=6$). Three species of invasive fish were also collected: common carp (*C. carpio*) ($n=74$), goldfish (*C. auratus*) ($n=5$) and eastern mosquitofish (*G. holbrooki*) ($n=15$). All fish caught were apparently healthy, with no signs of disease. Fish were caught using boat electrofishing, euthanized and dissected immediately upon capture. Tissue specimens (liver and gills) were placed in RNALater and stored in a portable -80°C freezer, then later in an -80°C freezer in the laboratory at Macquarie University, Sydney. Tissue selection was based on previous studies (Shi et al. 2018; Geoghegan et al. 2018; Geoghegan et al. 2021), which show that liver and gill tissue serve as a rich source of viruses. To facilitate virus discovery, multiple individuals (one to ten) were pooled according to species and the location in which they were captured (Supplementary Table S2 and Fig. S1).

2.3 Total RNA Extraction and Transcriptome Sequencing

Frozen samples of liver and gill tissue were processed together as a single extraction for each sample. The combined tissues were placed in $600\mu\text{l}$ of lysis buffer containing 0.5 per cent foaming reagent (Reagent DX, Qiagen) and 1 per cent of β -mercaptoethanol (Sigma-Aldrich). Submerged tissue samples were homogenized with TissueRuptor (Qiagen) for one minute at 5,000 rpm. To further homogenize tissue samples and remove tissue residues, the homogenate was centrifuged at full speed for three minutes. The homogenate was carefully removed and RNA from the clear supernatant was extracted using the RNeasy Plus Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. Extracted RNA was quantified using NanoDrop (ThermoFisher) and RNA from each species was

pooled corresponding to the site in which they were captured, resulting in thirty-six sample libraries (Supplementary Table S2). RNA libraries were constructed using the Truseq Total RNA Library Preparation Protocol (Illumina). To enhance viral discovery and reduce the presence of non-viral reads, host ribosomal RNA (rRNA) was depleted using the Ribo-Zero-Gold Kit (Illumina) and paired-end sequencing (150 bp) was performed on the NovaSeq 500 platform (Illumina). Sample library construction, rRNA depletion and RNA sequencing were performed at the Australian Genome Research Facility.

2.4 Virus discovery

Raw Illumina sequence reads (forward and reverse) were initially quality trimmed with Trimmomatic v.0.39 (Bolger, Lohse, and Usadel 2014) then assembled into contigs *de novo* using Trinity RNA-seq v.2.8.5 (Haas et al. 2013), with the default parameter settings. Assembled contigs were annotated and compared against the NCBI nucleotide (nt) and non-redundant protein (nr) databases with an e-value threshold of 1×10^{-5} using BLASTn and Diamond (BLASTX) (Buchfink, Xie, and Huson 2014). To initially distinguish between invertebrate and vertebrate-associated viruses, contigs that matched viral sequences were inspected using Geneious v.11.1.5 (Kearse et al. 2012) and translated into amino acid sequences. Amino acid sequences were then used as a single query in additional sequence comparisons against the NCBI nt and nr databases using BLAST algorithms. This method was also used to remove false-positives (e.g. host genes and endogenous viral elements) from our analyses. To help exclude instances of index hopping, viral sequences that were identified in multiple libraries were also

inspected using Geneious Prime (www.geneious.com) and amino acid pairwise alignments between viral sequences were performed with Multiple Alignment using Fast Fourier Transform (MAFFT) v.7.450 (Kato and Standley 2013), using the E-INS-i algorithm. Abundances of identical viral transcripts were then calculated (see below) and sequences that were present at frequency of less than one per cent of that of the number of reads present in the dominant library were excluded. To determine whether a virus was novel, we followed the broad criteria specified by The International Committee on Taxonomy of Viruses (<http://www.ictvonline.org/>).

2.5 Inferring the evolutionary history of novel viral sequences

To determine the evolutionary history of the viruses identified in this study and further distinguish between vertebrate and invertebrate-associated viruses (which are usually phylogenetically distinct), we estimated phylogenetic trees using amino acid sequences of stable genomic regions such as RNA-dependent RNA polymerase (RdRp) or DNA polymerase for DNA viruses. To this end, we combined our sequences with background sequences for each respective virus family taken from NCBI/GenBank. Amino acid sequences were aligned with MAFFT v.7.450 using the E-INS-i algorithm. To remove ambiguous regions in the sequence alignment, amino acid sequences were trimmed using trimAl v.1.2 (Capella-Gutiérrez, Silla-Martínez, and Gabaldón 2009). To estimate phylogenetic trees, selection of the best-fit model of amino acid substitution was determined using the Akaike information criterion (AIC), corrected AIC, and the Bayesian information criterion with the ModelFinder function (-m MFP) in IQ-TREE (Nguyen et al. 2015; Kalyaanamoorthy et al. 2017). Sequence data were analysed using a maximum likelihood (ML) approach in IQ-TREE, with 1,000 bootstrap replicates. Phylogenetic trees were annotated with FigTree v.1.4.2. and further edited using Adobe Illustrator (<https://www.adobe.com>).

2.6 Virome composition

To quantify the relative abundance of viral transcripts within the host transcriptome, the RNA-Seq by Expectation (RSEM) value was estimated using Trinity (Haas et al. 2013), and raw counts from each transcript were standardized against the total number of reads within the given sequencing library. We also used this approach to estimate the relative abundance of a host reference gene, ribosomal protein S13 (RPS13), which is stably expressed in fish. To assess any differences in virome composition between hosts and sites, we calculated alpha diversity (virome richness and Shannon diversity) using Rhea packages (Lagkouvardos et al. 2017). Generalized linear models (GLM) were used to identify the impact of host taxonomy (i.e. species), host geography (i.e. site), water temperature, water pH, water turbidity and species origin (i.e. invasive or native) on both vertebrate-associated virus composition (abundance, richness and diversity) and those viruses likely associated with non-fish hosts: the latter should not be affected by aspects of fish biology and hence effectively constitute a negative control. All GLM models were tested using a likelihood-ratio test (χ^2) and a Tukey's post hoc analysis (glht) was performed using the *multcomp* package (Hothorn, Bretz, and Westfall 2008). To assess viral diversity between samples, we calculated beta diversity using a Bray Curtis dissimilarity with the *phyloseq* package (McMurdle and Holmes 2013). Differences in virome

composition between native and invasive species were calculated using permanova (Adonis test), with the *vegan* package (Dixon 2003). All statistical analyses were carried out on RStudio V1.2.1335 and plotted using the *ggplot2* package (Valero-Mora 2015).

3. Results

We characterized the viromes of ten freshwater ray-finned fish species across seven taxonomic orders (two invasive and five native) at thirteen locations across the Murray–Darling Basin in Australia. Total RNA-sequencing was performed on thirty-six libraries, resulting in a median of 76,528,534 (range 66,015,138–95,168,951) reads per library. *De novo* assembly of the sequencing reads resulted in a median of 617,588 contigs (range 198,446–1,989,596) per library, with a total of 23,976,218 contigs generated. Analysis of the host reference gene, RPS13, revealed abundances of 0.000001–0.0002 per cent, suggesting an inconsistent sequence coverage across all RNA libraries, which may have impacted virus discovery (Fig. 2).

3.1 Abundance and diversity of viruses

We identified eighteen viral sequences that were associated with vertebrate hosts and a further twenty-seven that were likely associated with algae, invertebrates and protists in the freshwater environment (Supplementary Figs. S1 and S2). Because such non-vertebrate viruses were likely derived from diet or contamination of gill tissue, we primarily focused on vertebrate-associated viruses.

Among the likely vertebrate-associated viruses, we identified viral sequences from fourteen viral families. With the exception of a novel poxvirus (family *Poxviridae*), a double-stranded DNA virus, all the viruses identified possessed RNA genomes. The most abundant vertebrate-associated viral transcripts were those assigned to the *Arenaviridae* (49% of all vertebrate-associated viruses), *Hepeviridae* (20%), *Chuviridae* (21%), *Astroviridae* (3%) and *Flaviviridae* (2%) families. Other likely vertebrate viral transcripts detected were assigned to the *Coronaviridae* (<1%), *Caliciviridae* (<1%), *Picornaviridae* (<1%), *Paramyxoviridae* (<1%), *Hantaviridae* (<1%), *Bornaviridae* (<1%), *Poxviridae* (<1%), *Reoviridae* (<1%) and *Rhabdoviridae* (<1%) families. The most common vertebrate-associated viruses identified were astroviruses, detected in three host species (eastern mosquitofish, Murray–Darling rainbowfish, spangled perch). In addition, arenaviruses were detected in two host species (western carp-gudgeon, eastern mosquitofish) along with hepeviruses (common carp, eastern mosquitofish). All other viruses were identified in one host species (Fig. 3).

Among the viruses likely associated with non-vertebrate hosts (i.e. those infecting arthropods, fungi, plants and protozoans), a large proportion (70%) were unclassified, comprising picorna-like viruses, rhabdo-like viruses, tombus-like viruses and narna-like viruses (Shi et al. 2016) (Supplementary Fig. S1). We also detected viral transcripts that could be assigned to the *Nodaviridae* (27.1%), *Permutotetraviridae* (1.3%), *Dicistroviridae* (1.2%) and *Phenuiviridae* (<1%) families. Although viruses within the *Nodaviridae* have been shown to infect fish (Hameed et al. 2019), all of the nodavirus sequences identified here clustered with viruses from invertebrate hosts (Supplementary Fig. S2), strongly suggesting they were similarly associated with fish diet or contamination of gill tissue.

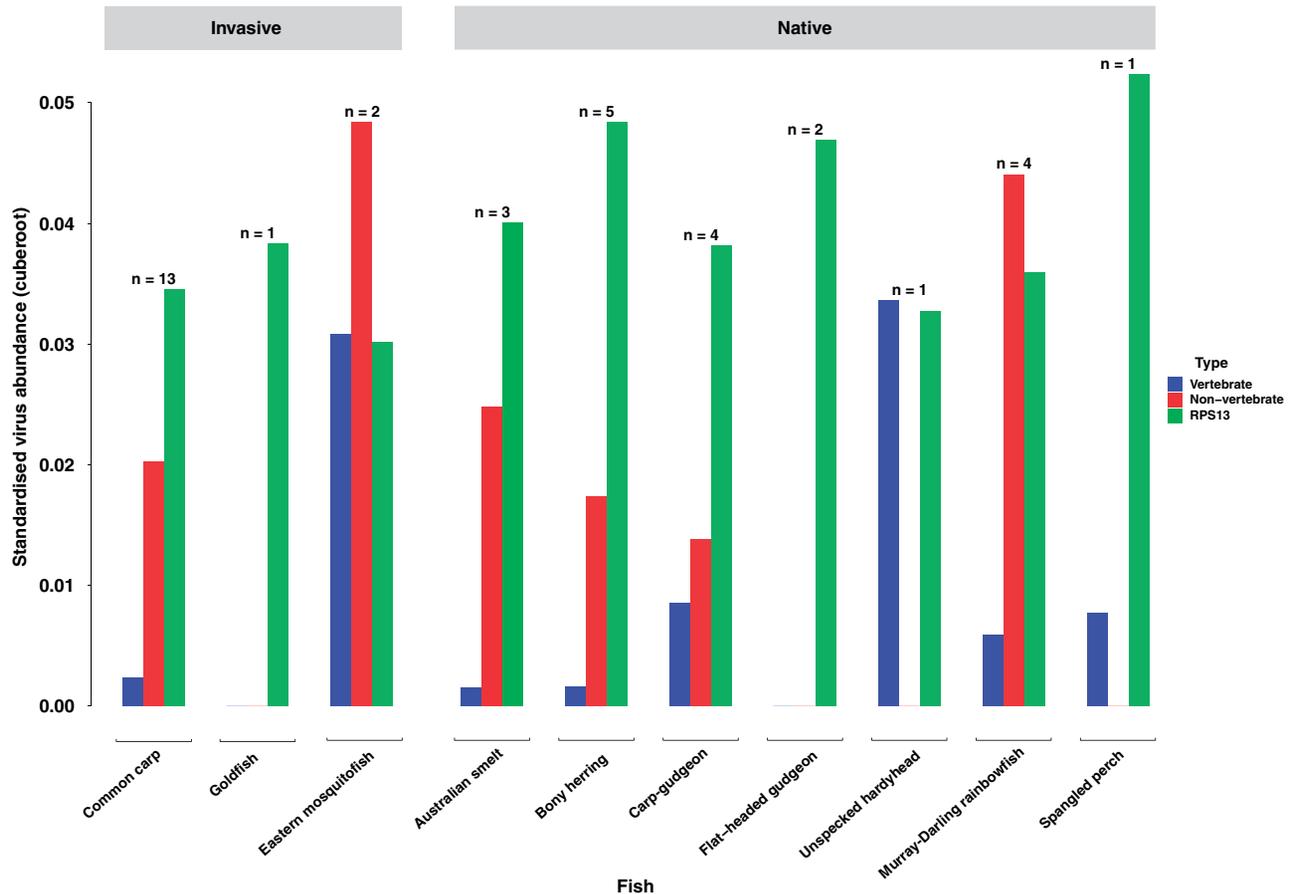


Figure 2. Mean standardized viral abundance across all libraries. Clustered bar chart reveals differences in viral abundance between invasive and native fish species. Blue bars represent vertebrate-associated viruses; red bars represent non-vertebrate-associated viruses; and green bars represent host reference gene RPS13. Number of sequencing libraries for each fish species is displayed above bars.

3.2 Phylogenetic relationships of vertebrate-associated viruses

To infer the phylogenetic relationships and hence the evolutionary history of the viruses newly identified here, we focused on stable genomic regions such as the RdRp in RNA viruses and DNA polymerase in the case of the novel poxvirus. Using these genomic regions, we identified seven negative-sense single-stranded RNA (-ssRNA) viruses (families *Arenaviridae*, *Bornaviridae*, *Chuviridae*, *Hantaviridae*, *Paramyxoviridae*, *Rhabdoviridae*), nine positive-sense single-stranded RNA (+ssRNA) viruses (families *Astroviridae*, *Caliciviridae*, *Coronaviridae*, *Flaviviridae*, *Hepeviridae*, *Picornaviridae*), one double-stranded RNA (dsRNA) virus (family *Reoviridae*) and one double-stranded DNA (dsDNA) virus (family *Poxviridae*). We now describe each of these groups in turn.

3.2.1 -ssRNA viruses

We identified -ssRNA viruses that occupied phylogenetic positions that were broadly indicative of long-term virus–host co-divergence, with many fish viruses falling basal to reptile, avian and mammalian viruses (Fig. 4). Notably, we identified two novel arenaviruses that clustered with members of the newly formed *Antennavirus* genus that includes fish hosts (Shi et al. 2018; Abudurexiti et al. 2019). *Western carp-gudgeon arenavirus* found at Narrabri creek shared 36.9 per cent amino acid sequence similarity with its closest relative, *Wenling frogfish arenavirus 1* (Shi et al. 2018) and grouped with recently discovered

arenaviruses from pygmy gobies (*Eviota zebrina*) sampled from an Australian coral reef (Geoghegan et al. 2021). *Eastern mosquitofish arenavirus* found in the Macquarie River shared 84.5 per cent amino acid sequence similarity with its closest available relative, *Wenling frogfish arenavirus 1*.

The divergent hantavirus detected in the Murray–Darling rainbowfish falls basal to mammalian hantaviruses (orthohantaviruses) and clustered with members of the *Actantavirinae* and *Agantavirinae* subfamilies that include ray-finned and jawless fish hosts (Shi et al. 2018; Abudurexiti et al. 2019) (Fig. 4). This virus had only 27.3 per cent amino acid similarity with its closest relative, *Bern perch virus* (NCBI/GenBank: QGM12349.1). Broad patterns of virus–host co-divergence can similarly be seen in the cultervirus identified in carp from Lake Burrendong. BLAST analysis identified *sharpbelly cultervirus* (Shi et al. 2018) as the closest relative of all genomic regions, including the L gene (93% amino acid similarity), glycoprotein (86.3%) and nucleoprotein (92.9%).

Our virological survey revealed the complete genome of a novel chuvirus in the unspecked hardyhead in the Edward River. This *Hardyhead chuvirus* displayed three open reading frames, representing the L protein (RdRp), glycoprotein and nucleoprotein. Our analysis identified *Guangdong red-banded snake chuvirus* (Shi et al. 2018) as the closest relative of the L protein (44% amino acid similarity), *Wenling fish chu-like virus* (Shi et al. 2018) as the closest relative of the glycoprotein (41%) and *Herr Frank virus 1* (Argenta et al. 2020) as the closest relative of the

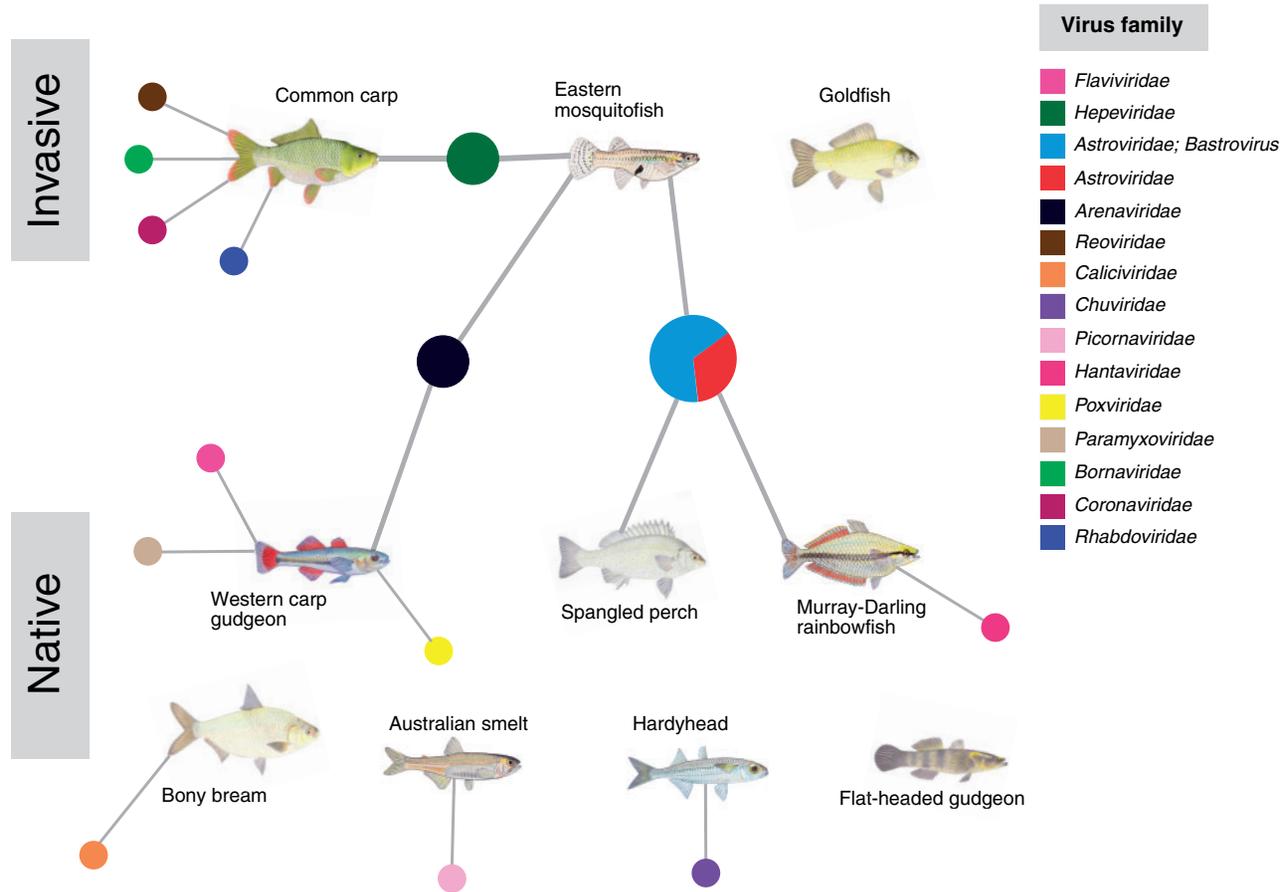


Figure 3. Network diagram displaying vertebrate-associated viruses identified in native and invasive freshwater fish. Colours of each node represent a virus family. Both goldfish and flat-headed gudgeon contained non-vertebrate-associated viruses.

nucleoprotein (34%). *Hardyhead* chuvirus formed a distinct phylogenetic clade with all other vertebrate-associated chuviruses (Fig. 5).

We also detected a novel paramyxovirus in western carp-gudgeon in the Bogan River. This divergent viral sequence shared 35.2 per cent L gene amino acid similarity with its closest relative, *Wenling tonguesole paramyxovirus* (genus *Cynoglossusvirus*, family *Paramyxoviridae*) (Shi et al. 2018). These viruses grouped with *Wenzhou pacific spadenose shark paramyxovirus* (genus *Scoliodonvirus*), together falling basal to other members of the *Paramyxoviridae* family. In addition, a novel rhabdovirus in common carp similarly formed a distinct clade, basal to other fish-infecting rhabdoviruses. This virus shared 35.7 per cent amino acid L gene sequence similarity with *Beihai dimarhabdovirus* that was also identified in fish (Shi et al. 2018) and clustered with other dimarhabdoviruses, including those found in the spotted paddle-tail newt from China (Shi et al. 2018) and the big brown bat (*Eptesicus fuscus*) from the USA (NCBI/Genbank: QPO14166.1). Across all genera within the *Rhabdoviridae*, lyssaviruses were the closest relatives to this clade (Fig. 4), with *Murray-Darling carp rhabdovirus* sharing 31.6 per cent amino acid L gene similarity with *rabies lyssavirus* (Tordo et al. 1988).

3.2.2 +ssRNA viruses

We identified a viral sequence in common carp that shared 50.7 per cent RdRp sequence similarity with *Pacific salmon nidovirus* (family *Coronaviridae*) (Mordecai et al. 2019). *Murray-Darling carp*

letovirus also exhibited sequence similarity (46.2%) with gamma-coronaviruses, including *bottlenose dolphin coronavirus* and *beluga whale coronavirus* (Mihindukulasuriya et al. 2008; Woo et al. 2014). This virus grouped with both *Pacific salmon nidovirus* and *Microhyla letovirus* (Bukhari et al. 2018), which together form an outgroup to all other coronaviruses (Fig. 6).

We also identified a novel flavivirus (genus *Flavivirus*, family *Flaviviridae*) in western carp-gudgeon in the Bogan River. This viral sequence exhibited thirty-three to thirty-six per cent NS5 amino acid sequence similarity with its closest relatives, *Cyclopterus lumpus virus* (Skoge et al. 2018), *Tamana bat virus* (de Lamballerie et al. 2002), *salmon flavivirus* (Soto et al. 2020) and *Wenzhou shark flavivirus* (Shi et al. 2018). All these viruses fall basal to vector-borne viruses within the genus *Flavivirus* (Fig. 7).

Among other positive-sense RNA viruses identified, a novel astrovirus, calicivirus and picornavirus all grouped with other fish hosts and expanded the phylogenetic diversity of these virus families (Fig. 7). The novel astrovirus identified in Murray-Darling rainbowfish shared forty per cent RdRp amino acid similarity with *Wuhan astro-like virus* (Shi et al. 2018). This virus clustered with other astro-like viruses discovered in fish, including *Beihai mudskipper astro-like virus* (Shi et al. 2018) and *Guangdong catfish astro-like virus* (Shi et al. 2018). This was similarly observed in *Australian smelt picornavirus*, which clustered with picornaviruses found in other freshwater fish, including those from eels (*Anguilla anguilla*) (Fichtner et al. 2013) and carp sampled from China (Shi et al. 2018). The *Caliciviridae* includes two genera that infect fish: saloviruses associated with

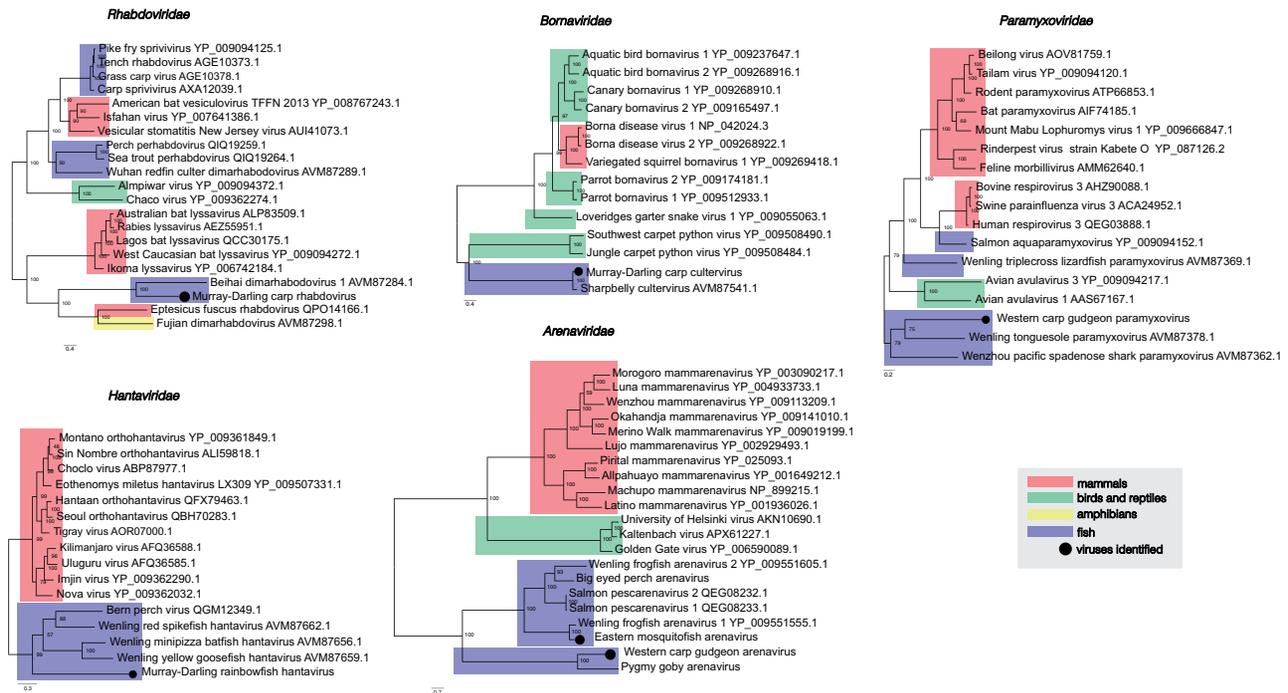


Figure 4. Phylogenetic relationships of negative-sense single-stranded vertebrate-associated viruses identified in this study. Viruses identified here are shown as a black circle. ML trees were estimated using amino acid sequences of the RdRp gene and mid-point rooted for clarity only. Bootstrap values are represented as a percentage with branches scaled to amino acid substitutions per site. Tree branches are highlighted to represent host class: red, mammals; green, birds and reptiles; yellow, amphibians; and blue, fish. Tip labels represent virus name and NCBI/GenBank accession numbers.

salmonid hosts and minoviruses associated with cyprinid hosts (Vinjé et al. 2019). Recently, several caliciviruses have been discovered in ray-finned and jawless fish (Mikalsen et al. 2014; Shi et al. 2018). The novel calicivirus identified in bony herring expands the diversity of fish viruses as it shared eighty per cent amino acid similarity with Atlantic salmon calicivirus (Mikalsen et al. 2014) and clustered with other freshwater fish caliciviruses, including *Wuhan carp calicivirus* and *Guangdong pseudoherculiter dispar calicivirus* (Shi et al. 2018).

3.2.3 dsRNA viruses

We identified a novel dsRNA virus in carp in the Castlereagh River that could be assigned to the *Reoviridae*. This divergent virus shared forty per cent RdRp amino acid similarity with its closest relative, *Wenling scaldfish reovirus* (Shi et al. 2018), together forming a clade basal to the genus *Aquareovirus* that are known to cause considerable disease in some fish species (Chen et al. 2018) (Fig. 8).

3.2.4 dsDNA viruses

A key observation of our study was the absence of Cyprinid herpesviruses, including CyHV-3, as well as other *Alloherpesviridae*, in any of the thirty-six RNA libraries. Similarly, although members of the *Hepadnaviridae* are commonly detected in fish (Dill et al. 2016; Lauber et al. 2017; Geoghegan et al. 2018; Geoghegan et al. 2021) they were notably absent in our samples. The only DNA virus detected in this study was a novel poxvirus (*Poxviridae*) identified in western carp-gudgeon. This virus shared DNA polymerase amino acid sequence similarity with *salmon gill poxvirus* (SGPV) (61%) (Gjessing et al. 2015). We also detected other genomic regions such as DNA-dependent RNA polymerase subunit rpo22 (49.5%), DNA-dependent RNA polymerase subunit rpo19 (40%), DNA binding virion core protein 11L

(28.1%), A16L (32.9%) and SGPV079 (40%) (Supplementary Table S1). Both SGPV and *western carp-gudgeon poxvirus* form a highly divergent clade within the subfamily *Chordopoxvirinae* that is strongly indicative of virus-host co-divergence (Fig. 9).

3.3 Virome composition, ecological and environmental factors

We next examined whether and how vertebrate virome composition in a range of native and introduced Murray–Darling Basin fish was associated host ecological factors, namely host species, geography (i.e. sampling site), water quality (temperature, pH and turbidity). GLMs revealed host species ($\chi^2 = 7.5 \cdot 10^{-6}$, $df = 9$, $P = 0.001$) as the best predictor of viral abundance (i.e. the standardized number of viral sequencing reads) (Fig. 2). In particular, the eastern mosquitofish had significantly higher viral abundance compared to Australian smelt (Tukey: $z = 3.976$, $P = 0.002$), bony herring (Tukey: $z = 4.334$, $P = 0.001$), western carp-gudgeon (Tukey: $z = 4.019$, $P = 0.002$), common carp (Tukey: $z = 4.665$, $P = 0.001$), flat-headed gudgeon (Tukey: $z = 3.632$, $P = 0.001$), goldfish (Tukey: $z = 3.632$, $P = 0.001$) and rainbowfish (Tukey: $z = 4.110$, $P = 0.001$). However, the high viral abundance in the eastern mosquitofish was driven by one sample containing an extremely high abundance of arenaviruses, accounting for seventy-six per cent of its total vertebrate virome and forty-nine per cent of all vertebrate-associated viral reads. We found no evidence for an association between viral abundance and host geography ($P = 0.111$), water turbidity ($P = 0.804$), water temperature ($P = 0.709$) nor water pH ($P = 0.141$).

We calculated alpha diversity to assess any differences in virome composition (abundance and diversity) between hosts and sites. This included the observed virus species richness (the number of viruses found in each sequencing library) and Shannon diversity (both the number of viral families and

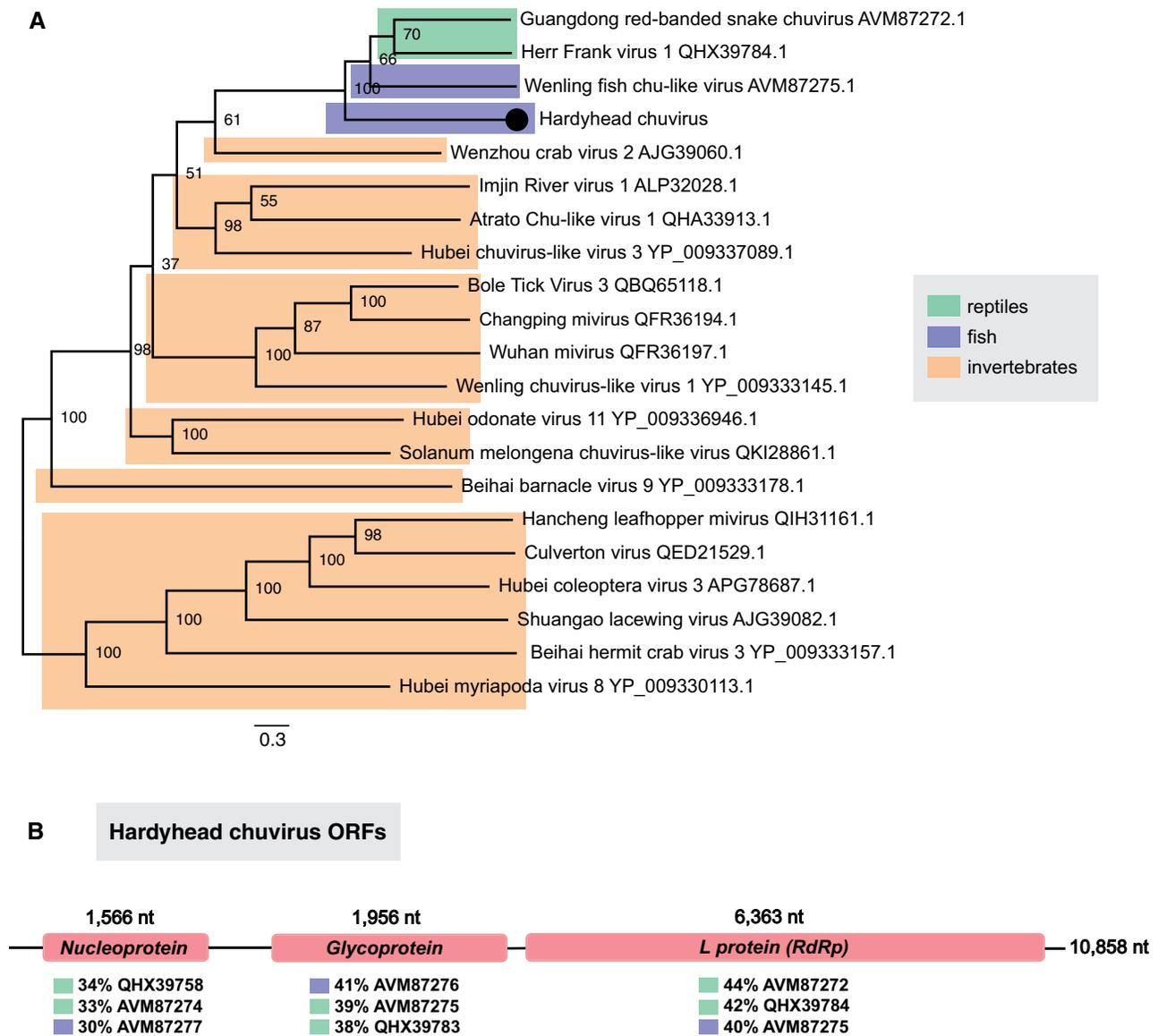


Figure 5. Phylogenetic relationships and genome organization of *hardyhead chuvirus*. (A) Phylogenetic relationships between viruses within the *Chuviridae*. Novel chuvirus identified in the unspiked *hardyhead* is represented as a black circle. The phylogenetic tree was estimated using amino acid sequences of the RdRp gene and midpoint rooted for clarity only. The scale bar represents amino acid substitutions per site. Bootstrap values are shown as a percentage. Tree branches are highlighted to distinguish between vertebrate and invertebrate-associated viruses: green, reptiles; blue, fish; and orange, invertebrates. Tip labels represent virus name and NCBI/GenBank accession numbers. (B) Genome organization of *hardyhead chuvirus*. Diagram illustrates the structure and length of each genomic segment. Percentages below each segment reveal the closest relatives (NCBI/GenBank accession number), with the L protein more related to reptile chuviruses; glycoprotein more related to fish chuviruses; and nucleoprotein more related to reptile chuviruses.

abundance of viral reads in a given host). We found no association between observed viral species richness and host species ($P = 0.286$), host geography ($P = 0.748$), water turbidity ($P = 0.826$), water temperature ($P = 0.625$) and water pH ($P = 0.115$). Similarly, there was also no observed association between Shannon diversity and host taxonomy ($P = 0.117$), host geography ($P = 0.551$), water turbidity ($P = 0.546$), water temperature ($P = 0.206$) and water pH ($P = 0.039$).

3.4 Virome composition of native versus invasive fish species

While carp and native fish species were sampled together at ten out of thirteen sites (Fig. 1), they shared no vertebrate-

associated viruses. Similarly, no vertebrate-associated viruses were shared among any of the native fish species. Although we identified hepeviruses in common carp and eastern mosquitofish, these were highly divergent and exhibited only twenty per cent amino acid similarity such that they reflect ancient divergence events. These viruses were also distinct in that *Murray-Darling carp hepevirus* clustered with ray-finned fish hosts (Shi et al. 2018), while *eastern mosquitofish hepevirus* formed a distinct basal clade with amphibian (Reuter et al. 2018), jawless fish and cartilaginous fish hosts (Shi et al. 2018) (Fig. 7).

While the native and invasive fish species largely had distinct viromes, two vertebrate-associated virus families were present in both: arenaviruses (western carp-gudgeon and

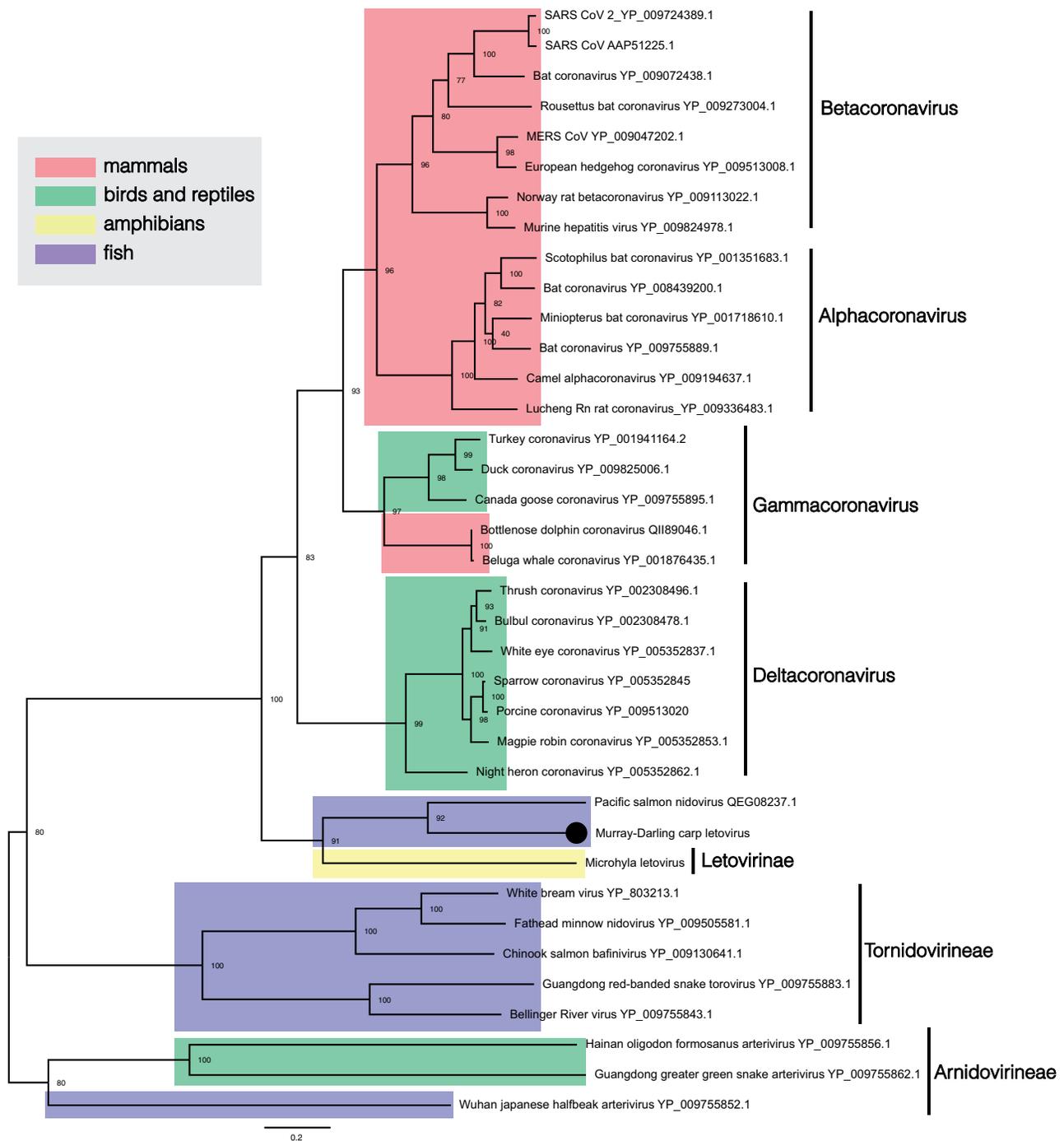


Figure 6. Phylogenetic relationships within the virus order Nidovirales. The novel letovirus detected in common carp is labelled with a black circle. The phylogenetic tree was estimated using the orf1ab polypeptide (polymerase associated) and midpoint rooted for clarity alone. Scale bar represents amino acid substitutions per site. Bootstrap values are shown as a percentage. Taxon labels are highlighted: blue, fish; red, mammals; green, birds and reptiles; yellow, amphibians. Virus taxonomic names are labelled to the right. Tip labels represent virus name and NCBI/GenBank IDs.

eastern mosquitofish) and bastroviruses (spangled perch and eastern mosquitofish) (Fig. 3). However, the arenaviruses identified in the eastern mosquitofish and western carp-gudgeon were highly divergent, exhibiting only 27.9 per cent amino acid similarity, with *western carp-gudgeon arenavirus* falling basal to *eastern mosquitofish arenavirus* (Fig. 4). The bastroviruses detected in eastern mosquitofish and spangled perch shared 57.1 per cent RdRp amino acid similarity and formed a distinct clade with other bastrovirus sequences identified in *Culex*

mosquitos (Sadeghi et al. 2018), bats (NCBI/GenBank: NC_035471.1) and sewage samples in Brazil (Dos Anjos, Nagata, and Melo 2017) (Supplementary Fig. S3). Because bastroviruses have genomic features that resemble hepeviruses (Reuter et al. 2018), both these contigs had matches to invertebrate and vertebrate-associated hepeviruses such that their true hosts could not be easily determined.

We also examined whether there were any differences in alpha and beta diversity between native and invasive

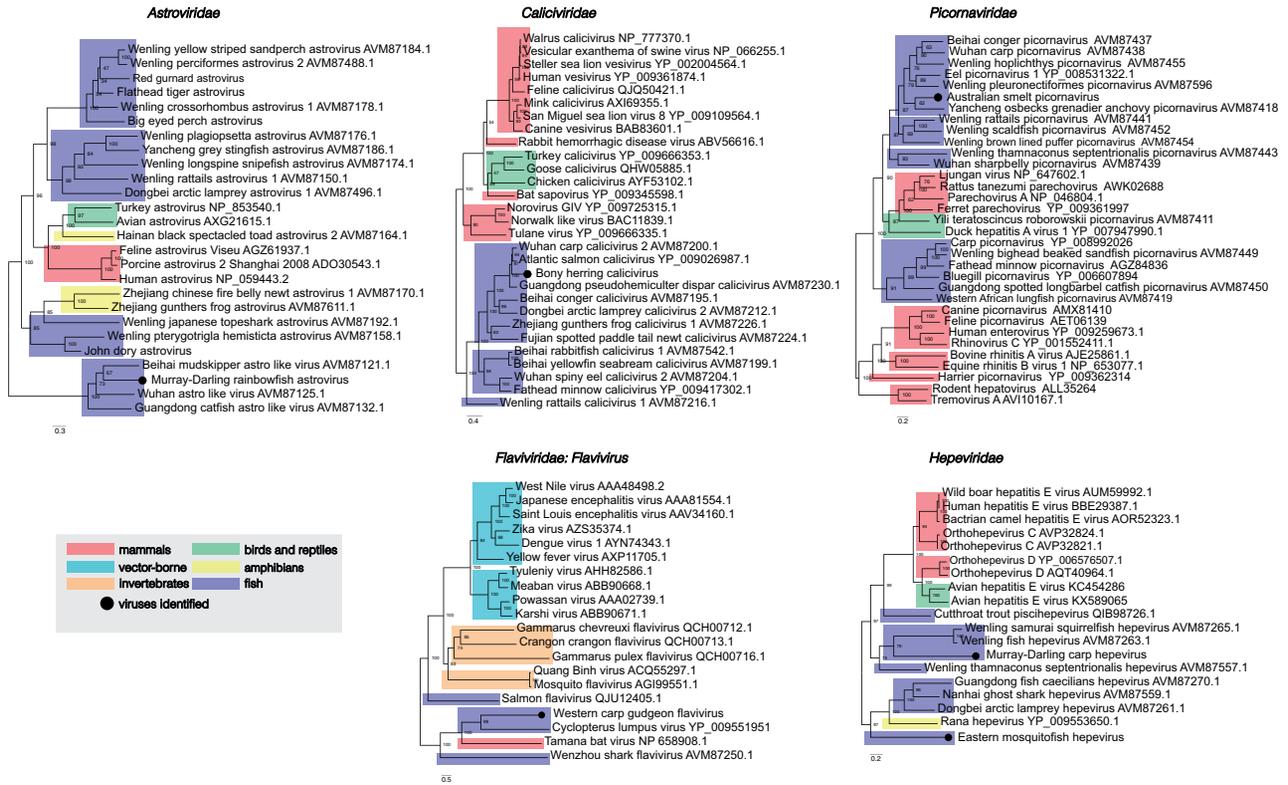


Figure 7. Phylogenetic relationships of positive-sense single-stranded vertebrate-associated viruses identified in this study. Viruses identified here are shown as a black circle. ML trees were estimated using amino acid sequences of the RdRp gene and NS5 gene for the novel flavivirus. Trees were mid-point rooted for clarity only and bootstrap values are represented as a percentage with branches scaled to amino acid substitutions per site. Tree branches are highlighted to represent host class: red, mammals; cyan, vector-borne viruses; green, birds and reptiles; yellow, amphibians; blue, fish; and orange, invertebrates. Tip labels represent virus name and NCBI/GenBank accession IDs.

freshwater fish. Accordingly, we found no association between host origin (i.e. invasive or native) and virome abundance ($P = 0.390$). When assessing alpha diversity, we similarly found no association between host origin and virome richness ($P = 0.626$) nor Shannon diversity ($P = 0.425$). Likewise, this result was also observed when examining beta diversity ($P = 0.602$).

3.5 Associations between host ecology and non-vertebrate viruses

To assess any associations between host ecology and non-vertebrate viruses, we similarly performed GLMs using the aforementioned ecological factors as a negative control. As expected, this revealed no association between non-vertebrate viral abundance and host taxonomy ($P = 0.200$), host geography ($P = 0.101$), host origin ($P = 0.998$), water turbidity ($P = 0.421$), water temperature (0.282) and water pH ($P = 0.343$). We similarly found no association between non-vertebrate virome richness and host taxonomy ($P = 0.204$), host geography ($P = 0.090$), host origin ($P = 0.675$), water turbidity ($P = 0.398$), water temperature ($P = 0.072$) and water pH ($P = 0.461$). We also found no evidence for an association between Shannon diversity and host taxonomy ($P = 0.691$), host geography ($P = 0.173$), host origin ($P = 0.876$), water turbidity ($P = 0.571$), water temperature ($P = 0.334$) and water pH ($P = 0.578$). This was also observed when assessing statistical associations between beta diversity and host species ($P = 0.684$), host origin ($P = 0.239$) and host geography ($P = 0.501$).

4. Discussion

Our meta-transcriptomic viral survey of native and invasive fish across the Murray–Darling Basin, Australia, revealed a high diversity and abundance of viruses, including the identification of 45 novel virus species that infected seemingly healthy fish or non-vertebrate hosts in the freshwater environment. Crucially, however, we observed no clear examples of recent cross-species transmission among any fish hosts, including between invasive and native species, nor any evidence for the presence of CyHV-3 from a total of thirty-six RNA sequencing libraries. Hence, these data provide further evidence of the absence of CyHV-3 in Australia (McCull and Crane 2013; McCull, Cooke, and Sunarto 2014; McCull et al. 2017). Similarly, our analysis failed to detect other cypriniviruses (i.e. CyHV-1, CyHV-2), despite previous reports of the presence of CyHV-2 in the Murray–Darling Basin (Stephens, Raidal, and Jones 2004; Becker et al. 2014).

While we observed no cypriniviruses in our analysis, such viruses typically induce latent infections in fish hosts and may become transcriptionally inactive (Michel et al. 2010). Thus, cypriniviruses in a latent stage may be difficult to detect using the transcriptome-based methods described here. Despite this, a viral survey of CyHV-1, CyHV-2 and CyHV-3 was previously conducted on carp across eight sites in the Murray–Darling Basin (McCull and Crane 2013). Using PCR amplicon sequencing, this study similarly failed to identify all three cypriniviruses in 849 carp DNA samples. This suggests that our findings may reflect a true absence of CyHV-3 in Australian waterways, although additional surveillance studies will be required to further address this issue.

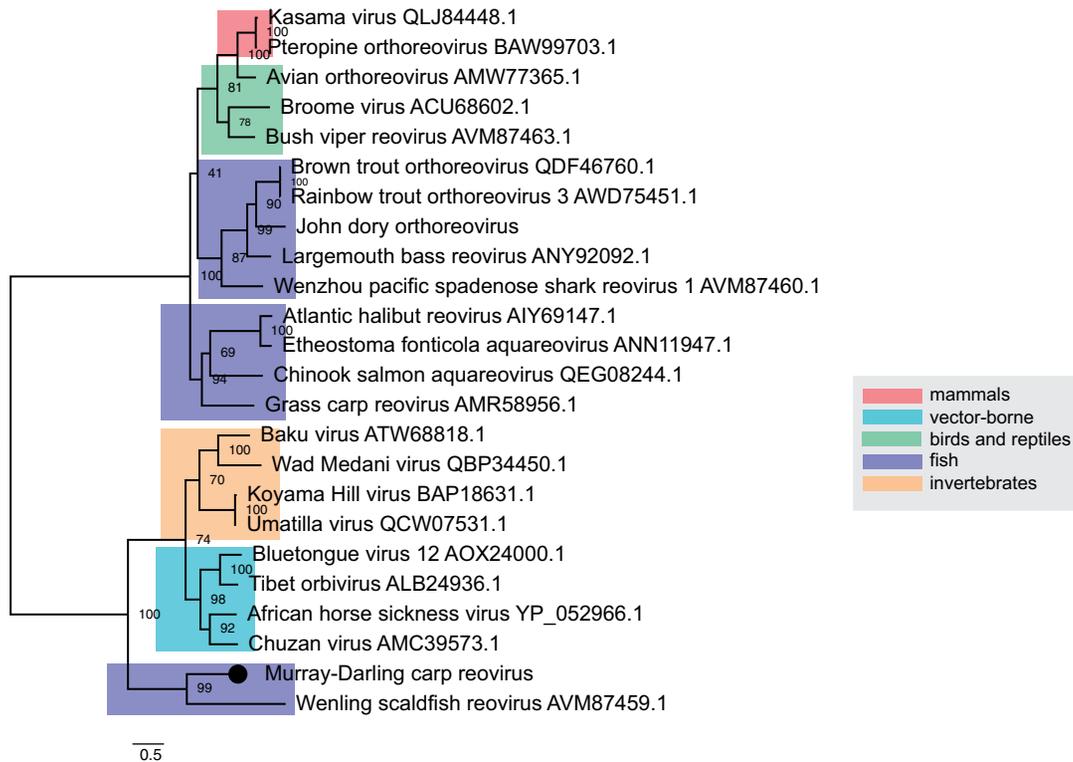


Figure 8. Phylogenetic relationships within the *Reoviridae*. Novel reovirus identified here shown as a black circle. ML tree was estimated using amino acid sequences of the RdRp gene. Tree was mid-point rooted for clarity only and bootstrap values are represented as a percentage with branches scaled to amino acid substitutions per site. Tree branches are highlighted to represent host class: red, mammals; cyan, vector-borne viruses; green, birds and reptiles; blue, fish; and orange, invertebrates. Tip labels represent virus name and NCBI/GenBank accession IDs.

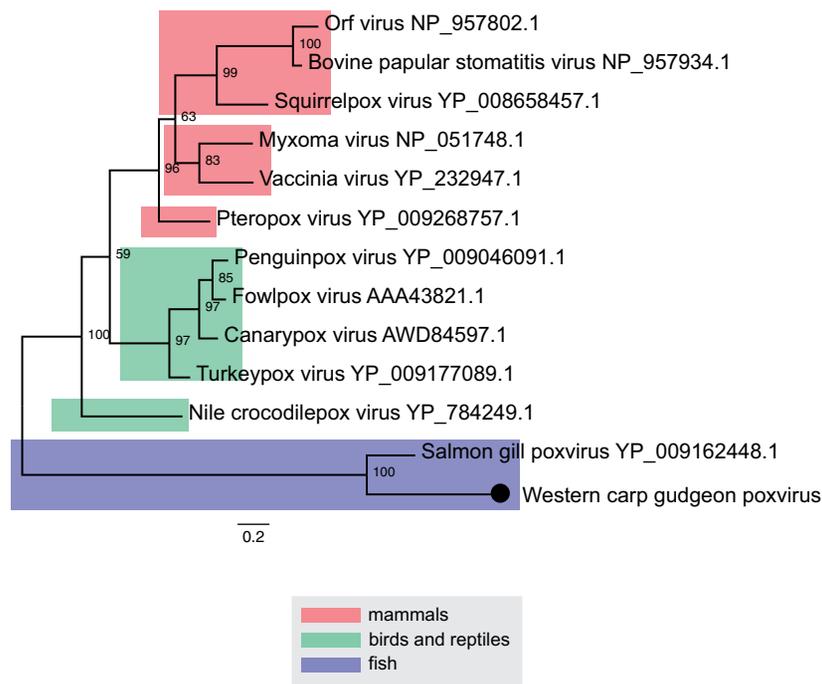


Figure 9. Phylogenetic relationships within the subfamily *Chordopoxvirinae* (family *Poxviridae*). Phylogenetic tree reveals virus-host co-divergence, with fish viruses falling basal to reptilian, avian and mammalian poxviruses. Novel poxvirus identified here shown as a black circle. ML tree was estimated using amino acid sequences of the DNA polymerase gene. Tree was mid-point rooted for clarity only and bootstrap values are represented as a percentage with branches scaled to amino acid substitutions per site. Tree branches are highlighted to represent host class: red, mammals; green, birds and reptiles; blue, fish. Tip labels represent virus name and NCBI/GenBank accession IDs.

Although carp and native fish have co-existed in the Murray–Darling Basin for over 50 years (Forsyth et al. 2013), they shared no vertebrate-associated viruses. The only instance of co-occurrence of viruses from the same family in both invasive and native species was the presence of arenaviruses in native carp-gudgeon and invasive mosquitofish. However, these viruses were so divergent that they likely represent ancient common ancestry rather than recent cross-species transmission (Fig. 4). Eastern mosquitofish, introduced into Australia during the early 1920s to control mosquito populations (Ayres, Pettigrove, and Hoffman 2010), are now widespread across the Murray–Darling Basin and have become a successful invasive species (Macdonald et al. 2012). Their abundance primarily impacts smaller native fish (such as carp-gudgeon, rainbowfish and hardyheads) since they typically outcompete these species and disrupt food webs (Macdonald et al. 2012). As well as being highly divergent, western carp-gudgeon arenavirus formed a basal clade with a recently discovered arenavirus in the pygmy goby sampled from an Australian coral reef (Geoghegan et al. 2021), a member of the same fish order (Gobiiformes). These data further suggest that arenaviruses may have been circulating in Gobiiform fishes (gobies and gudgeons) in Australia prior to the introduction of eastern mosquitofish.

We also identified a novel coronavirus—Murray–Darling carp letovirus—that shared 50.7 per cent amino acid sequence similarity with Pacific salmon nidovirus and 46.2 per cent amino acid similarity with gammacoronaviruses. The Coronaviridae (order Nidovirales) can be split into two subfamilies: the Orthocoronavirinae, associated with birds and mammals, and the Letovirinae associated with amphibians (Bukhari et al. 2018). Both Murray–Darling carp letovirus and Pacific salmon nidovirus formed a sister clade to only member of the Letovirinae subfamily, *Microhyla letovirus* (genus Alphaletovirus) identified in the ornamental pygmy frog (*Microhyla fissipes*) (Fig. 6), suggests that fish may be common and ancient hosts for the Letovirinae. It is also notable that Murray–Darling carp letovirus and Pacific salmon nidovirus are highly divergent from the other Nidovirales that are known to infect fish (e.g. Chinook salmon bafinivirus) (Cano et al. 2020).

The phylogenetic range of the Chuviridae largely incorporates invertebrate hosts with diverse genomes (segmented, unsegmented and circular) (Jun-Hua et al. 2015). Recently, chuviruses have been discovered in vertebrates, all possessing three segments (Shi et al. 2018; Argenta et al. 2020). The novel chuvirus detected here in the unspotted hardyhead displayed these genomic features with the L gene (RdRp), S gene (glycoprotein) and N gene (nucleoprotein) all related to fish and reptile viruses (Fig. 5). The phylogenetic position of this vertebrate clade suggests the ancestors of the viruses may be of invertebrate origin, particularly those that inhabit aquatic ecosystems. For instance, the closest related invertebrate viruses were Wenzhou crab virus (Jun-Hua et al. 2015), Imjin River virus (mosquitos) (Hang et al. 2016) and Atrato chu-like virus (mosquitos). Similarly, chuvirus endogenous viral elements have been detected in several freshwater fish species (Shi et al. 2018).

We identified a novel flavivirus in western carp-gudgeon from the Bogan River. This virus falls basal to mammalian vector-borne viruses in phylogenetic trees, grouping with viruses from other vertebrate hosts including *Cyclopterus lumpus virus*, *Tamana bat virus* and *Wenzhou shark flavivirus*. Although western carp-gudgeon flavivirus was detected in apparently healthy fish, in vivo flavivirus replication was recently demonstrated in Chinook salmon (*Oncorhynchus tshawytscha*) that were associated with abnormal mortalities in the Eel River, California (Soto

et al. 2020). While there is still no clear link between flavivirus infection, transmission and disease in aquatic hosts, these data suggest that flaviviruses may be common in fish species. Moreover, the basal phylogenetic positions of aquatic flaviviruses also suggest that these viruses may be the ancestors of notable vector-borne viruses (Fig. 7). Nevertheless, gaps still remain in the evolutionary history of the genus *Flavivirus* and will likely be bridged with additional metagenomic studies.

In broad terms, the evolutionary histories of many vertebrate viral families appear to generally follow patterns of long-term virus–host co-divergence, albeit with regular cross-species transmission (Shi et al. 2018; Zhang et al. 2018). This evolutionary pattern can be observed in the phylogenies of the culterivirus, poxvirus and arenaviruses identified here. The *Bornaviridae* contain three genera with eleven currently classified viral species that infect mammals, birds and reptiles (Amarasinghe et al. 2019). The only fish virus identified to date falls within the genus *Culterivirus*, comprising *Sharpbelly culterivirus* from China (Shi et al. 2018). We identified this virus (i.e. transcripts with 93% L gene amino acid similarity) in common carp in Australia. Intriguingly, both fish hosts are members of the *Cyprinidae* that date as far back as the Cretaceous to Jurassic periods (Hughes et al. 2018; Betancur et al. 2017). Molecular clock dating using endogenous viral elements also showed that culterviruses likely emerged early on during the course of vertebrate evolution, more than fifty million years ago (Shi et al. 2018).

Patterns of long-term virus–host co-divergence can also be seen in the evolutionary history of the *Chordopoxvirinae*. *Western carp-gudgeon poxvirus* expands the host range of the *Chordopoxvirinae* subfamily within the *Poxviridae*, forming a highly divergent clade with the only other fish-infecting chordopoxvirus discovered to date, SGPV (Fig. 9). Since its classification in 2015, several cases of SGPV have been identified in farmed salmon with complex gill disease, although the reservoir host is unknown (Gjessing et al. 2017). The phylogeny of the *Chordopoxvirinae* mirrors that of vertebrate hosts, strongly suggesting long-term virus–host co-divergence (Fig. 9). Similarly, the phylogeny of the *Arenaviridae* displays a basal fish clade that is characterized by long branches with a large degree of divergence (Fig. 4).

On this evolutionary backbone of ancient virus–host co-divergence, we also detected cases of cross-species virus transmission during evolutionary history, although the time-scales of these events are uncertain. For example, we discovered a novel reovirus that formed a basal divergent clade to other fish viruses within the genus *Aquareovirus* (Fig. 8). Murray–Darling carp reovirus was more closely related to viruses that infect scaldfish (Shi et al. 2018) rather than other cyprinid hosts, which are highly susceptible to reovirus infection (e.g. 80% mortality in grass carp) (Chen et al. 2018). These patterns were also observed in the phylogeny of the *Rhabdoviridae*, with Murray–Darling carp rhabdovirus forming a distinct phylogenetic clade with other recently discovered rhabdoviruses in fish, the big brown bat (NCBI/Genbank: QPO14166.1) and the spotted paddle-tail newt (Shi et al. 2018). Rhabdoviruses exhibit a very broad host range including invertebrates, plants, mammals, fish, amphibians, birds and reptiles (Bourhy et al. 2005). Notable among the *Rhabdoviridae* are the lyssaviruses that can cause high mortality in human populations (e.g. *rabies lyssavirus*). Intriguingly, Murray–Darling carp rhabdovirus and its closest relatives form a sister clade to the genus *Lyssavirus*, suggesting these viruses may have a fish-infecting ancestor (Fig. 4).

Although carp are widespread and abundant across the Murray–Darling Basin, they displayed lower viral abundance

than some of the other hosts sampled (Fig. 2). This could be partly explained by the large phylogenetic distance between carp and other fish in the Murray–Darling Basin. For instance, aside from bony herring, all the fish studied here are members of the *Acanthopterygii* (Percomorpha). This could also explain why invasive mosquitofish harboured similar viruses to native gudgeon species (e.g. arenaviruses). Although not always the case (Parrish et al. 2008), cross-species virus transmission often occurs between phylogenetically related hosts, particularly those that display conserved cell receptors (Longdon et al. 2014; Dennehy 2017). In addition, it has been widely suggested that introduced populations are associated with a lower pathogen prevalence and diversity than native species (Phillips et al. 2010; Mark et al. 2003). For example, because invasive species are often established from a small founder population, they likely carry and acquire only a small proportion of pathogens in the novel environment. Once a species rapidly becomes invasive, the diversity of pathogens in this population should remain small, such that the lack of disease likely facilitates the success of invasive species (Mark et al. 2003; Phillips et al. 2010).

It is important to note, however, that there were necessary variations within our sampling. For instance, carp and native fish species were sampled together at ten out of thirteen sites, with carp sampled from all thirteen sites (Fig. 1). In addition, all other fish species were sampled from one to five sites. While an artefact of the distribution of the fish, such gaps obviously limit the power of our statistical analyses and perhaps prevent the detection of ecological associations on virome composition within host species, including between invasive and native fish. In addition, due to animal ethics constraints, we were limited to only a subset of native fish species. Nevertheless, the native species examined in this study are generally those present in the highest densities.

Finally, our analysis detected no viruses that are listed as reportable notifiable aquatic diseases in the Murray–Darling Basin (Australian Health Committee 2020). Such notifiable aquatic diseases include epizootic haematopoietic necrosis virus (EHNV—*Iridoviridae*) and spring viraemia of carp virus (SVCV—*Rhabdoviridae*). EHNV is known to cause high-impact infections in redfin perch and is capable of infecting other freshwater fish in the Murray–Darling Basin, including silver perch (*Bidyanus bidyanus*), Macquarie perch (*Macquaria australasica*), Murray–Darling rainbowfish, freshwater catfish (*Tandanus tandanus*) and invasive mosquitofish (Becker et al. 2013). Although thought to be endemic to the Murray–Darling Basin (upper Murrumbidgee River), EHNV was last reported in 2012 (Becker et al. 2019). Similarly, we did not detect the emerging dwarf gourami iridovirus (*Iridoviridae*) that causes infectious spleen and kidney necrosis in several species of native Australian fish (Go et al. 2006; Rimmer et al. 2017).

In sum, our metagenomic surveillance revealed a marked lack of virus exchange between native and invasive fish species in the Murray–Darling Basin, including those viruses found in invasive common carp. At face value, these data suggest that there is minimal virus transmission from common carp to native fish species, although more extensive sampling is needed to fully address this issue. By investigating the viromes of native and invasive fish, we provide the first data on viruses that naturally circulate in a 2,200 km river system, enhancing our understanding of the evolutionary history of vertebrate viruses.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Acknowledgements

We thank Justin Stanger, Patrick Martin, Karen Scott, Michael Rodgers, Duncan McLay and Nick O'Brien from the Department of Primary Industries for their help with fish sampling. We also thank efishalbum.com for providing us with the fish images used in Figs. 1 and 3, which were used with permission.

Funding

This work was funded by a Recreational Fishing Trust grant (SF024) awarded to J.L.G., J.E.W. and E.C.H.; an ARC Discovery grant (DP200102351) awarded to E.C.H. and J.L.G.; a Macquarie University grant awarded to J.L.G.; and the Department of Biological Sciences at Macquarie University. E.C.H. is funded by ARC Australian Laureate Fellowship (FL170100022). J.L.G. is funded by a New Zealand Royal Society Rutherford Discovery Fellowship (RDF-20-UOO-007).

Data availability

All sequence reads generated in this project are available under the NCBI Short Read Archive (SRA) under BioProject PRJNA701716 and all consensus virus genetic sequences have been deposited in GenBank under accessions MW645018–MW645046.

Conflict of interest: None declared.

References

- Abudurexiti, A. et al. (2019) 'Taxonomy of the Order *Bunyavirales*: Update 2019', *Official Journal of the Virology Division of the International Union of Microbiological Societies*, 164: 1949–65.
- Amarasinghe, G. et al. (2019) 'Taxonomy of the Order *Mononegavirales*: Update 2019', *Official Journal of the Virology Division of the International Union of Microbiological Societies*, 164: 1967–80.
- Argenta, F. F. et al. (2020) 'Identification of Reptarenaviruses, Hartmanviruses, and a Novel Chuvirus in Captive Native Brazilian Boa Constrictors with Boid Inclusion Body Disease', *Journal of Virology*, 94: e00001–20.
- Australian Health Committee. (2020) Australia's National List of Reportable Diseases of Aquatic Animals, <<https://www.agriculture.gov.au/animal/aquatic/reporting/reportable-diseases#fnfish>> accessed 23 Feb 2021.
- Ayres, R., Pettigrove, V., and Hoffmann, A. (2010) 'Low Diversity and High Levels of Population Genetic Structuring in Introduced Eastern Mosquitofish (*Gambusia Holbrooki*) in the Greater Melbourne Area, Australia', *Biological Invasions*, 12: 3727–44.
- Barrett, J., Bamford, H., and Jackson, P. (2014) 'Management of Alien Fishes in the Murray–Darling Basin', *Ecological Management & Restoration*, 15: 51–6.
- Becker, J. A. et al. (2013) 'Experimental Infection of Australian Freshwater Fish with Epizootic Haematopoietic Necrosis Virus (EHNV)', *Journal of Aquatic Animal Health*, 25: 66–76.
- et al. (2014) 'Incursions of Cyprinid Herpesvirus 2 in Goldfish Populations in Australia despite Quarantine Practices', *Aquaculture*, 432: 53–9.
- et al. (2019) 'Geographic Distribution of (EHNV) in Freshwater Fish in South Eastern Australia: Lost Opportunity

- for a Notifiable Pathogen to Expand Its Geographic Range', *Viruses*, 11: 315.
- Bergmann, S. M. et al. (2010) 'Susceptibility of Koi × Crucian Carp and Koi × Goldfish Hybrids to Koi Herpesvirus (KHV) and the Development of KHV Disease (KHVD)', *Journal of Fish Diseases*, 33: 267–72.
- Betancur, R., R. et al. (2017) 'Phylogenetic Classification of Bony Fishes', *BMC Evolutionary Biology*, 17:
- Bolger, A. M., Lohse, M., and Usadel, B. (2014) 'Trimmomatic: A Flexible Trimmer for Illumina Sequence Data', *Bioinformatics*, 30: 2114–20.
- Bourhy, H. et al. (2005) 'Phylogenetic Relationships among Rhabdoviruses Inferred Using the L Polymerase Gene', *The Journal of General Virology*, 86(Pt 10): 2849–58.
- Buchfink, B., Xie, C., and Huson, D. H. (2014) 'Fast and Sensitive Protein Alignment Using DIAMOND', *Nature Methods*, 12: 59–60.
- Bukhari, K. et al. (2018) 'Description and Initial Characterization of Metatranscriptomic Nidovirus-like Genomes from the Proposed New Family *Abyssoviridae*, and from a Sister Group to the *Coronavirinae*, the Proposed Genus *Alphaletovirus*', *Virology*, 524: 160–71.
- Cano, I. et al. (2020) 'Isolation of a Chinook Salmon Bafinivirus (CSBV) in Imported Goldfish L. in the United Kingdom and Evaluation of Its Virulence in Resident Fish Species', *Viruses*, 12: 578.
- Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009) 'trimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses', *Bioinformatics*, 25: 1972–3.
- Chen, G. et al. (2018) 'Transcriptomics Sequencing Provides Insights into Understanding the Mechanism of Grass Carp Reovirus Infection', *International Journal of Molecular Sciences*, 19: 488.
- de Lamballerie, X. et al. (2002) 'Genome Sequence Analysis of Tamana Bat Virus and Its Relationship with the Genus *Flavivirus*', *The Journal of General Virology*, 83(Pt 10): 2443–54.
- Dennehy, J. J. (2017) 'Evolutionary Ecology of Virus Emergence', *Annals of the New York Academy of Sciences*, 1389: 124–46.
- Dill, J. A. et al. (2016) 'Distinct Viral Lineages from Fish and Amphibians Reveal the Complex Evolutionary History of Hepadnaviruses', *Journal of Virology*, 90: 7920–33.
- Dixon, P. (2003) 'VEGAN, a Package of R Functions for Community Ecology', *Journal of Vegetation Science*, 14: 927–30.
- Dos Anjos, K., Nagata, T., and Melo, F. L. d. (2017) 'Complete Genome Sequence of a Novel Bastrovirus Isolated from Raw Sewage', *Genome Announcements*, 5:
- Dudgeon, D. et al. (2006) 'Freshwater Biodiversity: Importance, Threats, Status and Conservation Challenges', *Biological Reviews*, 81: 163–82.
- Fichtner, D. et al. (2013) 'Characterization of a Novel Picornavirus Isolate from a Diseased European Eel (*Anguilla anguilla*)', *Journal of Virology*, 87: 10895–9.
- Forsyth, D. et al. (2013) 'Population Dynamics of Invading Freshwater Fish: Common Carp (*Cyprinus Carpio*) in the Murray-Darling Basin, Australia', *Biological Invasions*, 15: 341–54.
- Geoghegan, J. L. et al. (2018) 'Hidden Diversity and Evolution of Viruses in Market Fish', *Virus Evolution*, 4: vey031.
- et al. (2021) 'Virome Composition in Marine Fish Revealed by Meta-Transcriptomics', *Virus Evolution*, 7: veab005.
- Gjessing, M. C. et al. (2015) 'Salmon Gill Poxvirus, the Deepest Representative of the Chordopoxvirinae', *Journal of Virology*, 89: 9348–67.
- et al. (2017) 'Salmon Gill Poxvirus, a Recently Characterized Infectious Agent of Multifactorial Gill Disease in Freshwater- and Seawater-Reared Atlantic Salmon', *Journal of Fish Disease*, 40: 1253–65.
- Go, J. et al. (2006) 'The Molecular Epidemiology of Iridovirus in Murray Cod (*Maccullochella Pealii Pealii*) and Dwarf Gourami (*Colisa Lalia*) from Distant Biogeographical Regions Suggests a Link between Trade in Ornamental Fish and Emerging Iridoviral Diseases', *Molecular and Cellular Probes*, 20: 212–22.
- Haas, B. J. et al. (2013) 'De Novo Transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis', *Nature Protocols*, 8: 1494.
- Hang, J. et al. (2016) 'Genome Sequences of Five Arboviruses in Field-Captured Mosquitoes in a Unique Rural Environment of South Korea', *Genome Announcements*, 1: 4.
- Hayes, K., R. et al. (2014) 'Meeting the Challenge of Quantitative Risk Assessment for Genetic Control Techniques: A Framework and Some Methods Applied to the Common Carp (*Cyprinus Carpio*) in Australia', *Biological Invasions*, 16: 1273–88.
- Hothorn, T., Bretz, F., and Westfall, P. (2008) 'Simultaneous Inference in General Parametric Models', *Biometrical Journal*, 50: 346–63.
- Hughes, L., C. et al. (2018) 'Comprehensive Phylogeny of Ray-Finned Fishes (Actinopterygii) Based on Transcriptomic and Genomic Data', *Proceedings of the National Academy of Sciences*, 115: 6249–54.
- Iloize, M. et al. (2011) 'The Outbreak of Carp Disease Caused by CyHV-3 as a Model for New Emerging Viral Diseases in Aquaculture: A Review', *Ecological Research*, 26: 885–92.
- Johnson, P. T. J., and Paull, S. H. (2011) 'Ecology and Emergence of Diseases in Fresh Waters', *Freshwater Biology*, 56: 638–57.
- Jun-Hua, T. et al. (2015) 'Unprecedented Genomic Diversity of RNA Viruses in Arthropods Reveals the Ancestry of Negative-Sense RNA Viruses', *eLife*, 4: e05378.
- Kalyanamoorthy, S. et al. (2017) 'ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates', *Nature Methods*, 14: 587–9.
- Katoh, K., and Standley, D. M. (2013) 'MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability', *Molecular Biology and Evolution*, 30: 772–80.
- Kearse, M. et al. (2012) 'Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data', *Bioinformatics*, 28: 1647–9.
- Koehn, J. D. (2004) 'Carp (*Cyprinus carpio*) as a Powerful Invader in Australian Waterways', *Freshwater Biology*, 49: 882–94.
- Kopf, R. K. et al. (2019) 'Biocontrol in Australia: Can a Carp Herpesvirus (CyHV-3) Deliver Safe and Effective Ecological Restoration?', *Biological Invasions*, 21: 1857–70.
- Lagkouvardos, I. et al. (2017) 'Rhea: A Transparent and Modular R Pipeline for Microbial Profiling Based on 16S rRNA Gene Amplicons', *PeerJ*, 5: e2836.
- Lauber, C. et al. (2017) 'Deciphering the Origin and evolution of Hepatitis B Viruses by Means of a Family of Non-Enveloped Fish Viruses', *Cell Host & Microbe*, 22: 387–99.
- Longdon, B. et al. (2014) 'The Evolution and Genetics of Virus Host Shifts', *PLoS Pathogens*, 10: e1004395.
- Macdonald, J. I. et al. (2012) 'Do Invasive Eastern Gambusia Shape Wetland Fish Assemblage Structure in South-Eastern Australia?', *Marine and Freshwater Research*, 63: 659–71.
- Mark, E. T. et al. (2003) 'Introduced Species and Their Missing Parasites', *Nature*, 421: 628–30.
- Matsui, K. et al. (2008) 'Detection and Significance of Koi Herpesvirus (KHV) in Freshwater Environments', *Freshwater Biology*, 53: 1262–72.
- McCull, K. A., Cooke, B. D., and Sunarto, A. (2014) 'Viral Biocontrol of Invasive Vertebrates: Lessons from the past

- Applied to Cyprinid Herpesvirus-3 and Carp (*Cyprinus carpio*) Control in Australia', *Biological Control*, 72: 109–17.
- , and rane, M. S. J. (2013) *Cyprinid Herpesvirus 3 (CyHV-3): Its Potential as a Biological Control Agent for Carp in Australia*, PestSmart Toolkit Publication, Invasive Animals Cooperative Research Centre. https://pestsmart.org.au/wp-content/uploads/sites/3/2020/06/McCollCrane2012_KHV_ResearchReport.pdf
- et al. (2017) 'Cyprinid Herpesvirus 3 as a Potential Biological Control Agent for Carp (*Cyprinus carpio*) in Australia: Susceptibility of Non-Target Species', *Journal of Fish Diseases*, 40: 1141–53.
- McMurdie, P. J., and Holmes, S. (2013) 'Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data', *PLoS ONE*, 8: e61217.
- Michel, B. et al. (2010) 'Cyprinid Herpesvirus 3', *Emerging Infectious Diseases*, 16: 1835–43.
- Mihindukulasuriya, K., A. et al. (2008) 'Identification of a Novel Coronavirus from a Beluga Whale by Using a Panviral Microarray', *The Journal of Virology*, 82: 5084–8.
- Mikalsen, A., B. et al. (2014) 'Characterization of a Novel Calicivirus Causing Systemic Infection in Atlantic Salmon (*Salmo salar* L.): Proposal for a New Genus of Caliciviridae', *PLoS ONE*, 9: e107132.
- Mordecai, G. J. et al. (2019) 'Endangered Wild Salmon Infected by Newly Discovered Viruses', *eLife*, 8: e47615.
- Nguyen, L.-T. et al. (2015) 'IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies', *Molecular Biology and Evolution*, 32: 268–74.
- Parrish, C., R. et al. (2008) 'Cross-Species Virus Transmission and the Emergence of New Epidemic Diseases', *Microbiology and Molecular Biology Reviews*, 72: 457.
- Phillips, B., L. et al. (2010) 'Parasites and Pathogens Lag behind Their Host during Periods of Host Range Advance', *Ecology*, 91: 872–81.
- Reuter, G. et al. (2018) 'Detection of a Novel RNA Virus with Hepatitis E Virus-like Non-Structural Genome Organization in Amphibian, Agile Frog (*Rana dalmatina*) Tadpoles', *Infection, Genetics and Evolution*, 65: 112–6.
- Rimmer, A., E. et al. (2017) 'Susceptibility of a Number of Australian Freshwater Fishes to Dwarf Gourami Iridovirus (Infectious Spleen and Kidney Necrosis Virus)', *Journal of Fish Diseases*, 40: 293–310.
- Sadeghi, M. et al. (2018) 'Virome of >12 Thousand Culex Mosquitoes from throughout California', *Virology*, 523: 74–88.
- Sahul, H., A. S. et al. (2019) 'ICTV Virus Taxonomy Profile: Nodaviridae', *Journal of General Virology*, 100: 3–4.
- Shi, M. et al. (2018) 'The Evolutionary History of Vertebrate RNA Viruses', *Nature*, 561: 197–202.
- et al. (2016) 'Redefining the Invertebrate RNA Virosphere', *Nature*, 540: 539–43.
- , Zhang, Y.-Z., and Holmes, E. C. (2018) 'Meta-Transcriptomics and the Evolutionary Biology of RNA Viruses', *Virus Research*, 243: 83–90.
- Shimizu, T. et al. (2006) 'Survival of Koi Herpesvirus (KHV) in Environmental Water', *Fish Pathology*, 4: 153–7.
- Skoge, R. et al. (2018) 'New Virus of the Family Flaviviridae Detected in Lumpfish (*Cyclopterus Lumpus*)', *Official Journal of the Virology Division of the International Union of Microbiological Societies*, 163: 679–85.
- Soto, E. et al. (2020) 'First Isolation of a Novel Aquatic Flavivirus from Chinook Salmon (*Oncorhynchus Tshawytscha*) and Its Replication in a Piscine Animal Model', *Journal of Virology*, 94: e00337-20.
- Stephens, F., Raidal, Sr., and Jones, B. (2004) 'Haematopoietic Necrosis in a Goldfish (*Carassius auratus*) Associated with an Agent Morphologically Similar to Herpesvirus', *Australian Veterinary Journal*, 82: 167–9.
- Tolo, I. E. et al. (2021) 'Susceptibility of *Pimephales Promelas* and *Carassius auratus* to a Strain of Koi Herpesvirus Isolated from Wild *Cyprinus carpio* in North America', *Scientific Reports*, 11: 1985.
- Tordo, N. et al. (1988) 'Completion of the Rabies Virus Genome Sequence Determination: Highly Conserved Domains among the L (Polymerase) Proteins of Unsegmented Negative-Strand RNA Viruses', *Virology*, 165: 565–76.
- Valero-Mora, P. M. (2010) 'ggplot2: Elegant Graphics for Data Analysis', *Journal of Statistical Software*, 35:
- Vilizzi, L. et al. (2014) 'Ecological Effects of Common Carp (*Cyprinus carpio*) in a Semi-Arid Floodplain Wetland', *Marine and Freshwater Research*, 65: 802–17.
- Vinje, J. et al. (2019) 'ICTV Virus Taxonomy Profile: Caliciviridae', *The Journal of General Virology*, 100: 1469–70.
- Woo, P. C. Y. et al. (2014) 'Discovery of a Novel Bottlenose Dolphin Coronavirus Reveals a Distinct Species of Marine Mammal Coronavirus in Gammacoronavirus', *Journal of Virology*, 88: 1318–31.
- World Wildlife Fund. (2021) The World's Forgotten Fishes <https://www.panda.org/discover/our_focus/freshwater_practice/> accessed 23 Feb 2021.
- Zhang, Y.-Z., Shi, M., and., and Holmes, E. C. (2018) 'Using Metagenomics to Characterize an Expanding Virosphere', *Cell*, 172: 1168–72.
- et al. (2018) 'The Diversity, Evolution and Origins of Vertebrate RNA Viruses', *Current Opinion in Virology*, 31: 9–16.