VIRUS EVOLUTION

Virus Evolution, 2023, **9(1)**, 1–15
DOI: https://doi.org/10.1093/ve/vead011
Advance access publication date: 2 February 2023
Research Article

Limited cross-species virus transmission in a spatially restricted coral reef fish community

Vincenzo A. Costa,[1] David R. Bellwood,[2] Jonathon C.O. Mifsud,[1] Kate Van Brussel,[1] Jemma L. Geoghegan,[3,4,†] Edward C. Holmes,[1,‡] and Erin Harvey[1,*,§]

[1]Sydney Institute for Infectious Diseases, School of Medical Sciences, The University of Sydney, Sydney, NSW 2006, Australia, [2]Research Hub for Coral Reef Ecosystem Functions, College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia, [3]Department of Microbiology and Immunology, University of Otago, Dunedin 9016, New Zealand and [4]Institute of Environmental Science and Research, Wellington 5022, New Zealand
[†]https://orcid.org/0000-0003-0970-0153
[‡]https://orcid.org/0000-0001-9596-3552
[§]https://orcid.org/0000-0002-9379-2064
*Corresponding author: E-mail: erin.harvey@sydney.edu.au

Abstract

The Great Barrier Reef (GBR)—the largest coral reef ecosystem in the world—supports over 1,200 fish species with some of the highest population densities and diversities observed in vertebrates, offering a high potential for virus transmission among species. As such, the GBR represents an exceptional natural ecosystem to determine the impact of host community diversity on virus evolution and emergence. In recent decades, the GBR has also experienced significant threats of extinction, making it one of the most vulnerable ecosystems on the planet. Despite the global importance of the GBR, our understanding of virus diversity and connectivity in tropical reef fishes remains poor. Here, we employed metatranscriptomic sequencing to reveal the viromes of sixty-one reef fish species. This identified transcripts representing 132 putative viral sequences, 38 of which exhibited strong phylogenetic relationships with known vertebrate-associated viral genera, including a novel *Santee-Cooper ranavirus* (*Iridoviridae*). We found little evidence for virus transmission between fish species living within a very restricted geographical space—a 100-m^2 coral reef ecosystem—suggesting that there might be important host barriers to successful cross-species transmission despite regular exposure. We also identified differences in virome composition among reef fish families, such that cryptobenthic reef fishes—characterized by small body sizes and short life spans—exhibited greater virome richness compared to large reef fishes. This study suggests that there are important barriers to cross-species virus transmission and that successful emergence in a reef fish community likely requires active host adaptation, even among closely related host species.

Key words: virome; Great Barrier Reef; virus; metatranscriptomics; fish; virus discovery.

Introduction

Despite their long evolutionary history, extensive diversity, and complex ecological interactions, fish are greatly under-sampled in studies of viral ecology and evolution. Economically, fish contribute approximately 401 billion USD to the global economy and supply a yearly average of 20.5 kg per capita for consumption (FAO 2020). Yet fish face continual and potentially ruinous threats from emerging viral infections (Kibenge 2019). Climate-associated changes in species interactions are likely to have devastating ecological and economic consequences, particularly in the context of emerging infectious diseases (Burge et al. 2014). It was recently estimated that climate change may be responsible for the extinction of almost half of the economically important fish species in the tropical Pacific region by 2100 (Lam et al. 2020).

Tropical coral reefs are particularly vulnerable to biodiversity loss. In recent decades, global warming has caused numerous coral bleaching events worldwide, with detrimental cascading effects on reef ecosystem functioning and biodiversity (De'ath et al. 2012; Heron et al. 2016; Stuart-Smith et al. 2018). While tropical coral reefs make up only a small fraction of the marine environment, they support enormous biodiversity, accounting for approximately one-third of all currently described marine fish species (Spalding and Grenfell 1997). Many are 'cryptobenthic reef fishes', characterized by small body sizes (i.e. adult sizes of approximately 5 cm), short lifespans, cryptic behaviour, and benthic positioning on coral reefs (Depczynski and Bellwood 2003; Brandl et al. 2018). These diverse reef fish assemblages are of significant economic and cultural value to humans through aquaculture, fisheries, tourism, and the aquarium trade (Moberg and Folke 1999; Brandl et al. 2018; Woodhead et al. 2019).

Despite the economic and socioecological importance of reef fishes, little is known about the composition of their viromes or how ecological and phylogenetic variability within a reef fish community impacts cross-species virus transmission. Host community diversity likely plays a central role in virus emergence, as contact between donor and recipient hosts is a prerequisite

for virus transmission (Parrish et al. 2008). Yet, revealing the exact nature of that role has proven challenging (Ostfeld and Keesing 2012). A high diversity of hosts could provide more transmission opportunities for viruses, elevating the risk of disease emergence. Conversely, increased host community diversity may reduce the probability of disease emergence through the 'dilution effect', in which species richness provides more primary hosts for pathogens, in turn reducing disease occurrence in some other species (Schmidt and Ostfeld 2001; Ostfeld and Keesing 2012). As tropical coral reefs are biodiversity hotspots, they serve as an ideal natural ecosystem for exploring the impact of host diversity on the extent, pattern, and evolution of virus diversity. Indeed, reef fishes display some of the highest densities and highest diversities of potential vertebrate hosts on the planet. In our study location, a standard sampling area of 3.5 m^2 consistently supports between 50 and 150 fishes belonging to 15–25 species (Bellwood et al. 2006). These values equate to a mean density of approximately twenty-eight fishes per m^2. In terrestrial systems, vertebrates are typically at densities in the order of 0.001 per m^2, rarely, if ever, exceeding 1 per m^2 (Santini et al. 2018). Furthermore, these fish densities and richness are sustained year-round (Lefèvre et al. 2016), offering a continual potential for cross-species transmission in a readily transmitting aquatic medium.

Much of our knowledge on the tropical reef virosphere is skewed towards those viruses associated with coral species and their symbionts (Thurber et al. 2017). While epizootic infections have previously been reported in tropical reef fishes across the Western Atlantic and the Gulf of Mexico (Panek 2005), there are minimal data on viral diversity in fish from the Great Barrier Reef (GBR), Australia, although this is the largest and among the most threatened reef ecosystems in the world (Bellwood et al. 2004; De'ath et al. 2012; Stuart-Smith et al. 2018). A recent metatranscriptomic analysis of the pygmy goby (*Eviota zebrina*) identified three novel viruses (families *Arenaviridae*, *Hantaviridae*, and *Paramyxoviridae*) despite *Eviota* exhibiting maximum lifespans of between 60 and 100 days (including a 24–26-day pelagic larval phase) (Depczynski and Bellwood 2006; Geoghegan et al. 2021).

The rate of infectious disease emergence is expected to increase in marine environments, particularly in the context of climate change (Burge et al. 2014). As ocean temperatures continue to rise, many eastern Australian tropical fishes have begun to shift their distribution poleward to temperate reefs situated at higher latitudes through tropicalization (Vergés et al. 2014; Booth et al. 2018). While the broad-scale ecological impacts of tropicalization are becoming increasingly apparent in native temperate fishes (Vergés et al. 2014), it is unclear how this will impact virus ecology and infectious disease emergence.

Tropical ornamental fishes may also act as viral vectors of disease in economically important farmed species. For example, in Australia, dwarf gourami (*Trichopodus lalius*) can transmit infectious kidney and spleen necrosis virus (genus *Megalocytivirus* and *Iridoviridae*) to domestic fishes, often with detrimental impacts including disease outbreaks in iconic species such as Murray cod (*Maccullochella peelii*) (Lancaster, Williamson, and Schroen 2003; Go et al. 2006). Moreover, tropical wrasses (Labridae) have been considered effective biological control agents in aquaculture for their natural ability to consume pests (Barton et al. 2020). However, temperate cleaner wrasses have been reported to be important drivers of outbreaks of viral haemorrhagic septicaemia virus (*Rhabdoviridae*) in farmed salmonids (Murray 2016). As such, revealing viral diversity in ornamental tropical reef fishes is imperative to understanding the risk of disease emergence in both wild and domestic fish populations.

Revealing the nature of viral transmission within and among species in natural environments is central to understanding disease emergence. With exceptionally high species diversity and highly variable individual abundances and ecologies, coral reefs offer an opportunity to explore cross-species virus transmission in a complex natural high-diversity ecosystem. Our goal, therefore, was to employ total RNA sequencing (i.e. metatranscriptomics) to reveal the diversity, abundance, and composition of viruses infecting reef fishes from the GBR, utilizing a fish community from a 100-m^2 coral reef ecosystem. In particular, we aimed to (1) reveal viral diversity in tropical reef fishes, (2) determine the frequency of virus transmission within a spatially restricted reef fish community, (3) determine whether there are differences in virome composition between reef fish families, as well as between cryptobenthic reef fishes and large reef fishes, that differ in size, metabolic rate, lifespan, and fecundity (Brandl et al. 2018), and (4) identify novel viruses that may pose an emerging threat to Australian fisheries, aquaculture, and the aquarium trade.

Materials and methods

Animal ethics

Fish were collected under a Great Barrier Reef Marine Park Authority permit (G16/37684.1) and James Cook University Animal Ethics permit A2752.

Tropical reef fish sample collection

Fishes ($n = 192$) were collected in early April 2021 at Orpheus Island, GBR (18°36'44.3"S 146°28'59.4"E). All were "healthy" in that they were intact with good colour and no visible signs of disease, stress, or wasting. The vast majority were adult individuals. These included sixty-one species across sixteen reef fish families: Gobiidae (gobies) ($n = 30$ species), Labridae (wrasses) ($n = 6$), Pomacentridae (damselfishes) ($n = 5$), Blenniidae (blennies) ($n = 5$), Acanthuridae (surgeonfishes) ($n = 3$), Apogonidae (cardinalfishes) ($n = 2$), Monacanthidae (filefishes) ($n = 2$), Tetraodontidae (pufferfishes) ($n = 1$), Pseudochromidae (dottybacks) ($n = 1$), Chaetodontidae (butterflyfishes) ($n = 1$), Atherinidae (silversides, hardyheads) ($n = 1$), Serranidae (groupers) ($n = 1$), Tripterygiidae (triplefin blennies) ($n = 1$), Muraenidae (moray eels) ($n = 1$), Bythitidae (brotulas) ($n = 1$), and Ophichthidae (snake eels) ($n = 1$) (Supplementary Table S1). We sampled one to twelve individuals per species (±2.7 SD) (Supplementary Table S1). Of these species, forty-three (thirty-two cryptobenthic reef fishes and eleven large reef fishes) ($n = 148$ individuals) were collected from a reef fish community within a 100-m^2 sampling area (along the northern margin of Pioneer Bay). The other eighteen (seven cryptobenthic reef fishes and eleven large reef fishes) ($n = 44$ individuals) were collected from similar habitats (i.e. shallow fringing reefs) around Orpheus Island (Supplementary Table S1). Fishes within the focal sampling area were collected using an enclosed clove oil method, in which a small portion of the reef is enclosed within a fine net, and all fishes within the net were anaesthetized using clove oil (Ackerman and Bellwood 2002). Additional fish were collected using nets and/or dilute clove oil. All fish caught were placed either dissected (liver and gills) or whole in RNAlater and then transported to the lab on ice (Supplementary Table S1). Specimens were then stored at –80°C until RNA extraction.

RNA extraction, library preparation, and metagenomic next-generation sequencing

Tissue specimens (e.g. liver and gills or whole fish) were processed together as a single extraction for each individual fish sample.

The combined tissues were placed in 600 µl of lysis buffer containing 0.5 per cent foaming reagent (Reagent DX, Qiagen) and 1 per cent of β-mercaptoethanol (Sigma-Aldrich). Tissue samples were homogenized with a TissueRuptor (Qiagen) for up to 1 minute at 5,000 rpm. To remove tissue residues, the homogenate was centrifuged at full speed for 3 minutes. RNA from the clear supernatant was extracted using the RNeasy Plus Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol.

RNA was quantified using a UV–Vis cuvette spectrophotometer (DeNovix, Delaware, USA) and a parallel capillary electrophoresis instrument (Fragment Analyzer; Agilent, CA, USA). RNA from each individual fish was then pooled within each species (Supplementary Table S1), resulting in a total of sixty-one RNA sequencing libraries. These libraries were constructed using the Truseq Total RNA Library Preparation Protocol (Illumina). Host ribosomal RNA was depleted with the Ribo-Zero Plus Kit (Illumina), and paired-end sequencing (150 bp) was carried out on the NovaSeq 6000 platform (Illumina). To reduce the impact of index hopping and false virus–host assignments, each library was sequenced on two different lanes. Library construction and metatranscriptomic sequencing were performed by the Australian Genome Research Facility.

Virome assembly

Sequencing reads were first quality trimmed using Trimmomatic v.0.38 (Bolger, Lohse, and Usadel 2014) and then *de novo* assembled into contigs using MEGAHIT v.1.2.9, with default parameter settings (Li et al. 2015b). Assembled contigs for each library were used as queries against the NCBI nucleotide (nt) and non-redundant protein (nr) databases using BLASTn and Diamond (BLASTx) (Buchfink, Xie, and Huson 2015). To enable the identification of divergent viral sequences, we used an e-value search threshold of 1×10^{-5} (Shi et al. 2016, 2018; Geoghegan et al. 2021). All contigs with matches to viral sequences (NCBI/GenBank taxid: 10239) were predicted as open reading frames (ORFs) using Geneious Prime (v.2022.0) (Kearse et al. 2012) (www.geneious.com). To remove any false positives, each putative viral ORF was translated into amino acid sequences and used as a query to perform a second BLAST search against the NCBI nt and nr databases. From this second search, ORFs with matches to host genes (e.g. from fish or bacteria) were deemed as false positives and removed. ORFs with top hits to viruses from the initial search were selected and utilized for further analysis. Viral contig completion and contamination were determined using CheckV (Nayfach et al. 2021). The coverage was assessed by mapping using BOWTIE2 v 2.3.3.1 (Langmead and Salzberg 2012).

Virus taxonomic assignment and phylogenetic analysis

We used a phylogenetic approach to taxonomically assign our putative viral sequences and infer their evolutionary histories. First, we aligned our putative viral sequences with complete sequences of related viruses available on NCBI/GenBank (August 2021) using MAFFT v.7.450, with the E-INS-i algorithm (Katoh and Standley 2013). While some of these sequences were partial, they are adequate for phylogenetic analysis (Geoghegan et al. 2021). We used background sequences from the similarity search as well as additional sequences listed by the International Committee of Viral Taxonomy (ICTV) (https://talk.ictvonline.org) for each viral family. These included the conserved RNA-dependent RNA polymerase (RdRp) for RNA viruses and the DNA polymerase and major capsid protein for DNA viruses. The amino acid sequence alignment was trimmed using TrimAl v.1.2 to remove ambiguously aligned regions with a gap threshold of 0.9 and a variable conserve value (Capella-Gutierrez, Silla-Martinez, and Gabaldon 2009). The best-fit model of amino acid substitution was estimated with the 'ModelFinder Plus' (-m MFP) flag in IQ-TREE (Nguyen et al. 2015; Kalyaanamoorthy et al. 2017). Using these data, we estimated phylogenetic trees using a maximum likelihood approach with 1,000 bootstrap replicates using IQ-TREE. Trees were annotated using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software%20/figtree/).

Phylogenetic position and level of sequence similarity were used to determine whether a viral sequence was likely infecting fish (i.e. 'vertebrate-associated') or derived from diet or environment (i.e. 'non-vertebrate'): the latter often exhibits considerable genetic divergence and hence are phylogenetically distinct (Shi et al. 2018; Zhang et al. 2018; Costa et al. 2021; Geoghegan et al. 2021). Vertebrate-associated viral sequences were classified as novel species according to similarity thresholds for each viral family as specified by the ICTV.

Viral genome annotation

Viral genomes were annotated using the NCBI conserved domain (CDD) search tool and InterProScan with the TIGRFAMs (v15.0), SFLD (v4.0), PANTHER (v15.0), SuperFamily (v1.75), PROSITE (v2022_01), CDD (v3.18), Pfam (v34.0), SMART (v7.1), PRINTS (v42.0), and CATH-Gene3D databases (v4.3.0) (Jones et al. 2014).

Removal of endogenous viral elements

To identify whether our putative viral contigs were expressed endogenous viral elements, we initially searched for disrupted ORFs using Geneious Prime (v.2022.0). Next, we screened for flanking host regions using CheckV and BLASTn. We also searched for flanking long terminal repeats as these are associated with integrated proviruses (Johnson 2019). Contigs with intact ORFs that were not flanked by host genes or repeats were subsequently used as a query against all vertebrate genomes available on NCBI using tblastn, with an e-value threshold of 1×10^{-20} (Shi et al. 2018). We used all published vertebrate genomes as no genomes are currently available for the reef fish studied here. Viral sequences identified as endogenous were removed from subsequent analyses.

RT-PCR validation of vertebrate-associated viruses

We used RT-PCR to validate the presence of the putative exogenous vertebrate-associated viruses detected here. Primers for each virus were designed using Primer3 v.0.4.0 (Untergasser et al. 2012). SuperScript IV One-Step RT-PCR (Invitrogen) was used to amplify viral genes from the total RNA of all sequencing library pools, with annealing temperatures ranging from 50°C to 55°C (Supplementary Table S2).

Abundance estimation and virome statistical analysis

We employed RNA-seq by Expectation-Maximization (v1.2.28) to quantify the relative abundance of transcripts within each fish species transcriptome (Haas et al. 2013). These included both viral genes and the stably expressed host reference gene, ribosomal protein S13 (RPS13), which was used to assess sequencing depth across libraries (Geoghegan et al. 2018, 2021). Abundance measures were standardized by dividing values against the total reads for each library. We calculated both alpha and beta diver-

sity to compare virome composition between reef fish families, as well as between cryptobenthic reef fishes ($n = 39$ species) and large reef fishes ($n = 22$ species). We also analysed the composition of the non-vertebrate virome as a form of internal control, as these viruses are not impacted by aspects of fish biology. Accordingly, we used Rhea scripts to calculate alpha diversity, including viral abundance, observed virome richness, and Shannon diversity (Lagkouvardos et al. 2017). Statistical comparisons of alpha diversity were modelled using generalized linear models and tested using a likelihood-ratio test (χ^2) and Tukey's post hoc analysis with the *multcomp* package (Hothorn, Bretz, and Westfall 2008). To compare viral communities between reef fish assemblages, we calculated beta diversity using a Bray–Curtis distance matrix with the *phyloseq* package (McMurdie and Holmes 2013). These data were then tested using permutational multivariate analysis of variance (permanova) with the *vegan* package (adonis) (Dixon 2003). All plots were constructed using *ggplot2* (Valero-Mora 2010).

Results

Total diversity and abundance of viruses discovered in tropical reef fish

We sequenced sixty-one metatranscriptomes of tropical reef fish species for virus discovery, resulting in a median of 99,658,615 (range 45,583,181–121,905,422) reads and 371,944 (range 161,447–626,467) contigs with an average N50 of 1,049 bp per library (Supplementary Table S1). From these data, we identified transcripts representing thirty-eight putative exogenous viral sequences that exhibited strong phylogenetic relationships with viral families and genera known to infect vertebrate species, comprising eleven negative-sense single-stranded RNA viruses (-ssRNA) (*Hantaviridae*, *Chuviridae*, *Orthomyxoviridae*, *Rhabdoviridae*, and *Paramyxoviridae*), eighteen positive-sense single-stranded RNA viruses (+ssRNA) (*Astroviridae*, *Hepeviridae*, *Picornaviridae*, *Flaviviridae*, and *Coronaviridae*), one double-stranded RNA virus (dsRNA) (*Reoviridae*), six single-stranded DNA viruses (ssDNA) (*Parvoviridae* and *Circoviridae*), and two double-stranded DNA viruses (dsDNA) (*Iridoviridae*) (Table 1; Fig. 1). Using PCR, we confirmed the presence of thirty-six of the thirty-eight vertebrate-associated viruses from our reef fish RNA libraries (Table 1). Two libraries contained an insufficient volume of RNA for virus detection. In total, thirty-two libraries (seventeen cryptobenthic reef fishes and fifteen large reef fishes) contained no vertebrate-associated viral sequences.

Our metatranscriptomic analysis also identified ninety-four viral sequences that shared sequence similarity and phylogenetically clustered with viruses associated with invertebrates (arthropods, crustaceans, molluscs, platyhelminthes, myriapods, and nematodes) and fungal hosts (i.e. the 'non-vertebrate' group; Supplementary Figs S1–S4). These viruses were classified within the *Flaviviridae* (27 per cent of total non-vertebrate virome abundance), *Narnaviridae* (24 per cent), *Picornaviridae* (20 per cent), *Nodaviridae* (10 per cent), *Hepeviridae* (7 per cent), *Solemoviridae*

Table 1. Description of vertebrate-associated viruses identified in this study.

Virus name	Reads	Virus family	Host	ORFs/ segments	Length (nt)	Complete genome	Novel	PCR
Goby astrovirus 2	24,100	*Astroviridae*	*Cabillus tongarevae*	3	7,226	Y	Y	+
Goby astrovirus 2	2,767	*Astroviridae*	*Istigobius decoratus*	3	7,085	Y	Y	+
Goby astrovirus 2	9,023	*Astroviridae*	*Istigobius nigroocellatus*	3	7,155	Y	Y	+
Goby astrovirus 2	1,184	*Astroviridae*	*Asterropteryx semipunctatus*	3	7,045	Y	Y	+
Goby astrovirus 1	2,863	*Astroviridae*	*Istigobius goldmanni*	3	7,087	Y	Y	+
Goby astrovirus 3	6,709	*Astroviridae*	*Luposicya lupus*	2	7,355	Y	Y	+
Blenniella astrovirus	59	*Astroviridae*	*Blenniella sp.*	1	744	N	Y	+
Salarias guttatus piscichuvirus	40,334	*Chuviridae*	*S. guttatus*	3	10,801	Y	Y	+
Abudefduf bengalensis circovirus	54	*Circoviridae*	*A. bengalensis*	1	873	N	Y	+
Enneapterygius tutuilae letovirus	26	*Coronaviridae*	*E. tutuilae*	1	856	N	Y	+
Chaetodon aureofaciatus hepacivirus	1,029	*Flaviviridae*	*C. aureofaciatus*	1	8,826	Y	Y	+
Luposicya lupus actinovirus	6,144	*Hantaviridae*	*L. lupus*	3	11,533	Y	Y	+
Pygmy goby hantavirus	972	*Hantaviridae*	*Eviota zebrina*	3	5,894	N	N	+
Trimma fangi actinovirus	106	*Hantaviridae*	*T. fangi*	2	1,319	N	Y	+
Pleurosicya actinovirus	123	*Hantaviridae*	*Pleurosicya* sp.	2	2,103	N	Y	+
Cephalopholis boenak actinovirus	9	*Hantaviridae*	*C. boenak*	1	1,029	N	Y	+
Istigobius decoratus hepevirus	16	*Hepeviridae*	*I. decoratus*	1	393	N	Y	+
Halichoeres melanurus ranavirus	9,835	*Iridoviridae*	*H. melanurus*	9	10,821	N	N	+
Enneapterygius tutuilae iridovirus	35,545	*Iridoviridae*	*E. tutuilae*	9	20,970	N	Y	+

(continued)

Table 1. (Continued)

Virus name	Reads	Virus family	Host	ORFs/ segments	Length (nt)	Complete genome	Novel	PCR
Cheilodipterus quinquelineatus orthomyxovirus	979	*Orthomyxoviridae*	*C. quinquelineatus*	5	9,090	N	Y	+
Pygmy goby paramyxovirus	238	*Paramyxoviridae*	*E. zebrina*	1	1,314	N	N	+
Enneapterygius tutuilae paramyxovirus	322	*Paramyxoviridae*	*E. tutuilae*	1	618	N	Y	+
Istigobius ornatus parvovirus	1,746	*Parvoviridae*	*I. ornatus*	3	4,284	Y	Y	+
Istigobius rigillus parvovirus	266	*Parvoviridae*	*I. rigillus*	1	1,656	N	Y	+
Ecsenius stictus parvovirus	2,707	*Parvoviridae*	*E. stictus*	2	4,286	Y	Y	+
Dinematichthys parvovirus	122	*Parvoviridae*	*Dinematichthys* sp.	1	1,236	N	Y	+
Luposicya lupus ichthama-parvovirus	12,328	*Parvoviridae*	*L. lupus*	1	1,299	N	Y	+
Blenniella picornavirus	2	*Picornaviridae*	*Blenniella* sp.	1	347	N	Y	+
Enneapterygius tutuilae picornavirus	225	*Picornaviridae*	*E. tutuilae*	1	3,840	N	Y	+
Exyrias picornavirus	44	*Picornaviridae*	*Exyrias* sp.	1	1,373	N	Y	+
Halichoeres marginatus picornavirus	27,506	*Picornaviridae*	*H. marginatus*	1	8,344	Y	Y	+
Neopomacentrus bankieri picornavirus	2,955	*Picornaviridae*	*N. bankieri*	1	8,408	Y	Y	+
Asterropteryx spinosa picornavirus	8,307	*Picornaviridae*	*A. spinosa*	1	8,405	Y	Y	+
Asterropteryx semipunctatus picornavirus	210	*Picornaviridae*	*A. semipunctatus*	1	992	N	Y	+
Pomacentrus moluccensis picornavirus	12	*Picornaviridae*	*P. moluccensis*	1	321	N	Y	+
Trimma okinawae rhabdovirus	13	*Rhabdoviridae*	*T. okinawae*	2	945	N	Y	+
Trimma striatum rhabdovirus	2	*Rhabdoviridae*	*T. striatum*	1	488	N	Y	-
Istigobius nigroocellatus reovirus	282	*Reoviridae*	*I. nigroocellatus*	6	6,966	N	Y	-

(4 per cent), *Totiviridae* (1 per cent), *Partitiviridae*, *Iflaviridae*, *Picobirnaviridae*, *Phenuiviridae*, *Qinviridae*, *Rhabdoviridae*, *Tombusviridae*, *Mimiviridae*, *Natareviridae*, and *Jingchuvirales* (all less than 1 per cent) (Li et al. 2015a; Shi et al. 2016). We also identified viruses related to negeviruses (2 per cent of non-vertebrate virus abundance) (Nunes et al. 2017) and quenyaviruses (<1 per cent) (Obbard et al. 2020) (Supplementary Fig. S1). While many of these families include viruses that are known to infect fishes (e.g. *Flaviviridae*, *Picornaviridae*, *Nodaviridae*, *Hepeviridae*, *Totiviridae*, and *Rhabdoviridae*), the viruses identified here were highly divergent and clustered with those that infect a broad range of invertebrate and fungal species (Supplementary Figs S1–S4). We therefore focused our analysis on vertebrate-associated viruses.

Genome organization and phylogenetic relationships of vertebrate-associated viruses

-ssRNA viruses

Five actinoviruses were identified (subfamily *Actantavirinae* and *Hantaviridae*) (Fig. 2E). All were novel, with the exception of pygmy goby hantavirus previously identified in *E. zebrina* (Geoghegan et al. 2021). We identified complete and partial actinovirus genomes in *Trimma fangi*, *Luposicya lupus*, *Cephalopholis boenak*, and *Pleurosicya* sp. (Fig. 2E; Supplementary Table S3). The genome of *L. lupus*actinovirus contained all three expected actinovirus genomic regions (Fig. 2E). The S segment possessed an additional ORF upstream of the nucleoprotein in antisense orientation, also seen in the genome of perch actinovirus, but not in other actinoviruses (Hierweger et al. 2021).

Two partial genomes of paramyxoviruses were identified, one in *E. zebrina* (i.e. pygmy goby paramyxovirus) (Geoghegan et al. 2021) and the other in *Enneapterygius tutuilae*. This virus was highly divergent from other classified ray-finned fish paramyxoviruses, including the genus *Aquaparamyxovirus* (Fig. 2A). We also identified partial genomes (L protein and nucleocapsid sequences) of two novel rhabdoviruses in *Trimma okinawae* and *Trimma striatum*. The *T. okinawae* rhabdovirus fell within the monophyletic *Tupavirus* group that infects birds and mammals (Kurz et al. 1986; Allison et al. 2011), whereas *T. striatum* rhabdovirus clustered with the

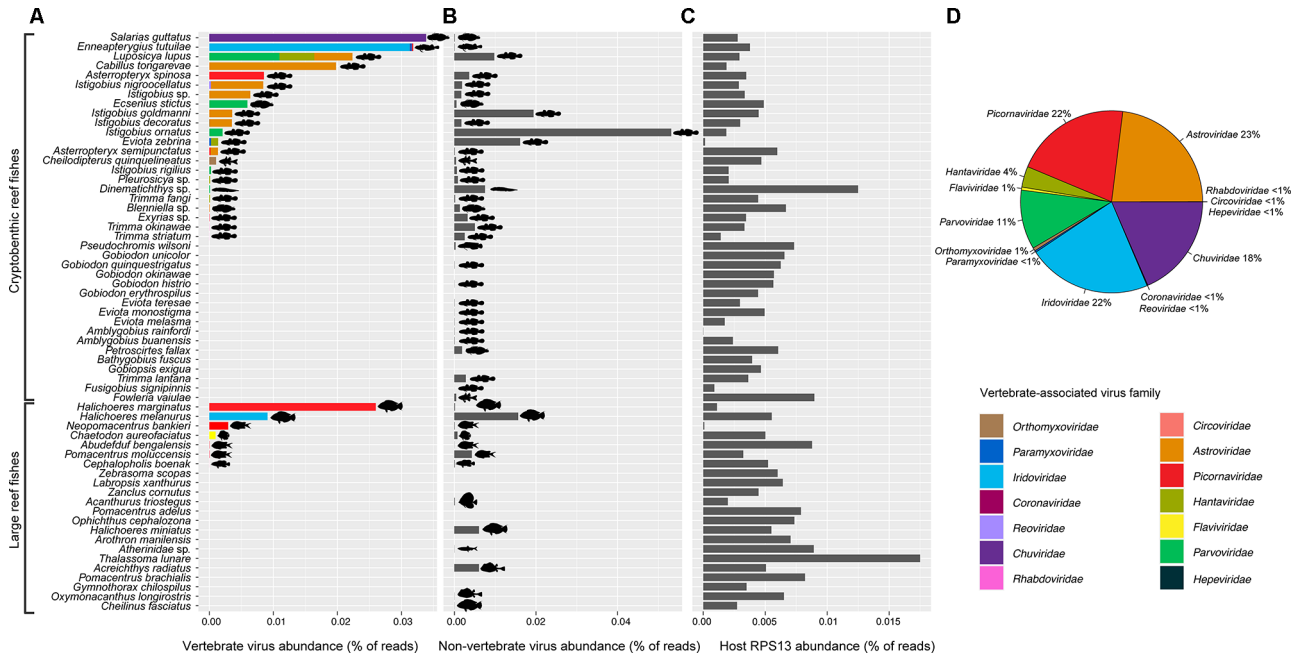

Figure 1. The standardized abundance of viral and host reads across reef fish libraries. The abundance (i.e. percentage of reads) of (A) likely vertebrate-associated viruses; (B) non–vertebrate-associated viruses (i.e. those from algae, fungi, coral, arthropods, crustaceans, and protists); and (C) host reference gene RPS13. Silhouettes represent viruses associated with fish species. (D) The total standardized abundance of vertebrate-associated viral families.

dimarhabdoviruses (Shi et al. 2018) (Fig. 2B). Among other -ssRNA viruses, we identified a novel pischichuvirus in *Salarias guttatus* and orthomyxovirus in *Cheilodipterus quinquelineatus* that showed sequence similarity to quaranjaviruses (Fig. 2D–E; Supplementary Table S3).

+ssRNA viruses

Astroviruses were detected in seven fish species. All were members of the Gobiidae, except for *Blenniella* sp. (Blenniidae). The astroviruses identified here clustered with those that infect a broad range of ray-finned and jawless fish species (Fig. 3D). Similarly, we identified a divergent virus from this group in *L. lupus*, tentatively named goby astrovirus 3. This virus had a genome of 7,355 nt with a predicted 5′ untranslated region (UTR) of 347 nt, two main ORFs, and a 3′ UTR of 347 nt. One ORF encoded the astrovirus ORF1a protein (protease), while the other encoded a polyprotein containing the RdRp and capsid protein. This arrangement is also seen in the genome of Wenling rattails astrovirus 5 but not in other currently described fish astroviruses that exhibit three main ORFs (Shi et al. 2018).

The complete genomes of three novel picornaviruses were identified in *Halichoeres marginatus, Neopomacentrus bankieri*, and *Asterropteryx spinosa*, and partial sequences of five novel picornaviruses were identified in *Asterropteryx semipunctatus*, *Pomacentrus moluccensis, E. tutuilae, Exyrias*, and *Blenniella* sp. (Fig. 3A). Although these fish species were members of the same community, all eight viruses were highly divergent. Phylogenetic analysis of RdRp revealed clustering with other fish picornaviruses, with the exception of *E. tutuilae* picornavirus and *Exyrias* picornavirus that clustered with *Ampivirus* identified in the common newt (*Lissotriton vulgaris*) (Reuter et al. 2015) (Fig. 3A). We also identified the full genome of a novel hepacivirus (*Flaviviridae*) in *Chaetodon aureofaciatus* with conserved genomic regions—Core C, Envelope E1, NS2, NS3, NS4, and NS5B—that fell within a distinct group of aquatic hepaciviruses (Shi et al. 2018) (Fig. 3B).

dsRNA viruses

The near-complete genome of a novel reovirus was identified in *Istigobius nigroocellatus*. This included six segments encoding the VP1 (guanyl transferase), VP2 (RdRp), VP3 (helicase), VP4 (NTPase), VP5 (outer capsid protein), and VP6 (core protein) proteins. Phylogenetic analysis of the VP2 gene revealed clustering within the genus *Aquareovirus*, sharing 48–60 per cent amino acid similarity with its closest relatives (Supplementary Table S3; Supplementary Fig. S5).

ssDNA viruses

We discovered a basal group of four parvoviruses that fell within the subfamily *Parvovirinae* (*Parvoviridae*) (Fig. 4A). The genome of *Ecsenius stictus* parvovirus was a single contig of 4,286 nt, containing two main ORFs. The left ORF was 1,374 nt and encoded the conserved non-structural protein NS1 that has DNA helicase and ATPase function (Pénzes et al. 2020). The right ORF encoded the structural VP1 protein (2,571 bp). *Istigobius ornatus* parvovirus had a genome of 4,284 nt and contained three main ORFs, one encoding NS1 (1,878 nt) and the other two encoding putative structural proteins (Fig. 4B). We also identified the partial genome of a novel ichthamaparvovirus in *L. lupus* and a novel circovirus in *Abudefduf bengalensis* that clustered within a distinct clade of fish circoviruses (Fig. 4C).

dsDNA viruses

The analysis of the *Halichoeres melanurus* meta-transcriptome identified nine conserved proteins sharing 99–100 per cent similarity with all currently described *Santee-Cooper ranavirus* isolates,

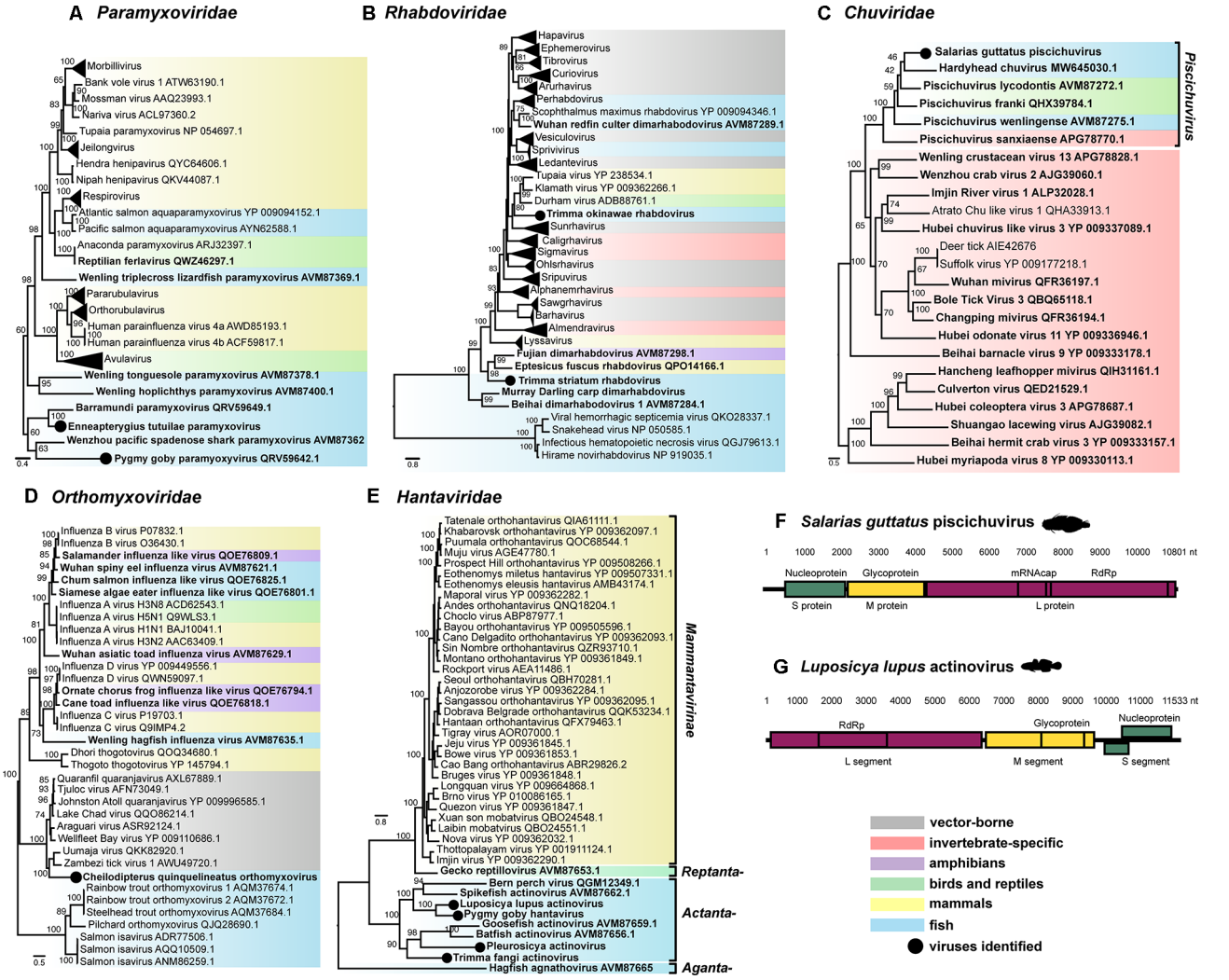

Figure 2. The genome structure and maximum likelihood phylogenies of the RdRp gene of -ssRNA viruses identified in this study. (A) The phylogeny of novel and related virus species of the *Paramyxoviridae*. (B) The phylogeny of the *Rhabdoviridae*. (C) The phylogeny of the *Chuviridae*. The genus *Piscichuvirus* is labelled to illustrate vertebrate-associated viruses. (D) The phylogeny of the *Orthomyxoviridae*. (E) The phylogeny of the *Hantaviridae*. The subfamilies *Mammantavirinae*, *Reptantavirinae*, *Actantavirinae*, and *Agantavirinae* are labelled to illustrate virus–host co-divergence. (F) The schematic genome of *Salarias guttatus* piscichuvirus. (G). The schematic genome of *Luposicya lupus* actinovirus. Trees were midpoint rooted for clarity. The scale bar represents the number of amino acid substitutions per site. Tip labels represent virus name with NCBI/GenBank accession. Viruses in the bold text represent sequences identified through metagenomics. Tree branches are highlighted to broadly represent host taxonomy. Schematic genome diagrams illustrate genome orientation, length, predicted ORFs, and gene products.

such as largemouth bass virus, mandarin fish ranavirus, koi ranavirus, doctor fish virus, and guppy virus 6 (*Iridoviridae*; Supplementary Table S3). We used the highly conserved major capsid protein and DNA polymerase for phylogenetic analysis (Whittington, Becker, and Dennis 2010), which further confirmed a novel *Santee-Cooper ranavirus* isolate, tentatively named *H. melanurus* ranavirus (Fig. 5). Largemouth bass virus and mandarin fish ranavirus are highly lethal in farmed populations, causing 95–100 per cent mortality (Zhang et al. 2017; Zhao et al. 2020). Our metatranscriptomic analysis also identified a novel iridovirus in *E. tutuilae* that fell sister to erythrocytic necrosis virus (ENV) (Pagowski et al. 2019) and clustered with other erythrocytic-like viruses identified in amphibians and reptiles (Wellehan et al. 2008; Alves de Matos et al. 2010; Grosset et al. 2014; Russo et al. 2021). ENV was the closest relative across all proteins identified (Supplementary Table S3).

Cross-species virus transmission in a reef fish community

Despite the large number of viruses identified, we only found evidence for one cross-species transmission within our GBR ecosystem. This involved astroviruses found in five different fish species that clustered together and exhibited high levels of amino acid sequence similarity (Fig. 3D). Specifically, phylogenetic comparisons of ORF1b (RdRp) revealed two viral species: goby astrovirus 1, identified in *Istigobius goldmanni*, and goby astrovirus 2, identified in *I. nigroocellatus*, *Istigobius decoratus*, *A. semipunctatus*, and *Cabillus tongarevae* (Fig. 3D). These two viruses exhibit 82.5 per cent amino acid sequence similarity across the virus genome, while the four sequences of goby astrovirus 2 had 96.8 per cent amino acid similarity (Fig. 3D).

We also detected related viruses (i.e. those from the same virus family) in several different fish species, including the *Hantaviridae*,

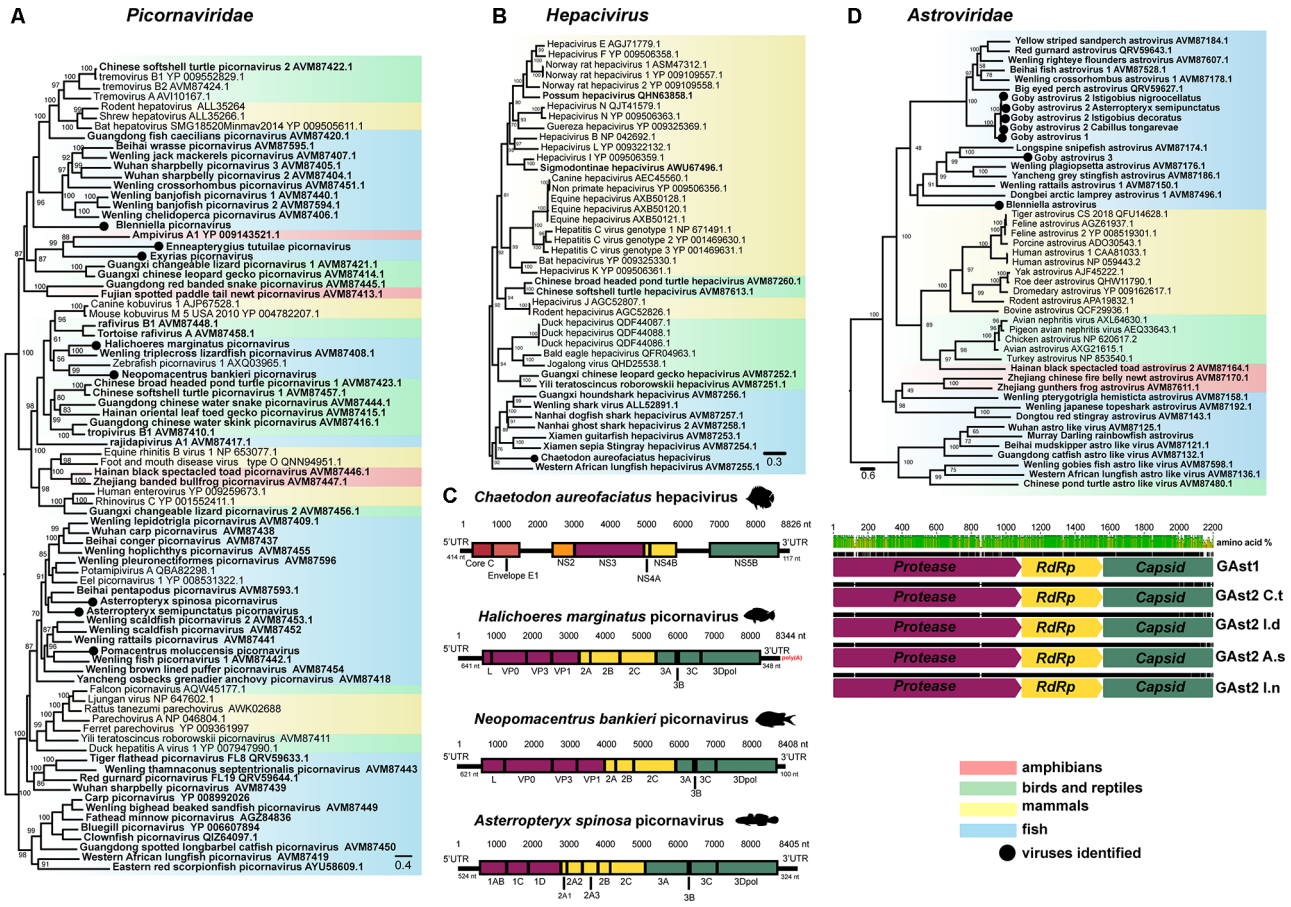

Figure 3. The genome organization and phylogenies of the RdRp gene of +ssRNA viruses identified in this study. (A) The maximum likelihood phylogeny of the *Picornaviridae*. (B) The phylogeny of the genus *Hepacivirus* (*Flaviviridae*). (C) Schematic genomes illustrate genome orientation, length, and gene products. (D) The phylogeny of the *Astroviridae* and amino acid alignment of goby astrovirus 2 (GAst2) isolates against goby astrovirus 1 (GAst1). GAst2 are labelled to represent host species: GAst2 C.t = *C. tongaravae*, I.d = *I. decoratus*, A.s = *A. semipunctatus*, I.n = *I. nigroocellatus*. Trees were midpoint rooted for clarity. The scale bar represents the number of amino acid substitutions per site. Tip labels represent virus name with NCBI/GenBank accession. Viruses in the bold text represent sequences identified through metagenomics. Trees are highlighted to broadly illustrate host taxonomy.

Rhabdoviridae, *Paramyxoviridae*, and *Picornaviridae* (Fig. 6A). However, most of these were sufficiently divergent in sequence that they likely reflect common ancestry rather than direct host jumping in the reef ecosystem. For example, of the eight picornaviruses identified in this study, the closest relatives were *A. spinosa* picornavirus and *A. semipunctatus* picornavirus that shared only 44 per cent amino acid similarity.

In addition, we identified viral species not directly infecting the fish themselves but rather associated with the local environment, diet, or microbiome (i.e. non-vertebrate viruses) that were also transmitted among reef fish assemblages. These were as follows: quenyaviruses (95 per cent amino acid similarity between *P. moluccensis* and *A. bengalensis*), flavi-like viruses (91.2 per cent between *I. goldmanni* and *I. ornatus*), narnaviruses (97.7 per cent between *I. goldmanni*, *I. nigroocellatus*, and *Istigobius rigillus*) and totiviruses (93.5 per cent between *Amblygobius buanesis* and *Amblygobius rainfordi*) (Supplementary Figs S1, S2, and S4; Fig. 6C). That there were more instances of cross-species transmission of non-vertebrate viruses compared to those viruses that actively replicate in fish suggests that the latter group are subject to strong host barriers, even among closely related species.

Comparisons of viral alpha and beta diversity between reef fish families

We compared vertebrate virome composition between reef fish families, as well as between cryptobenthic reef fishes and large reef fishes (i.e. that differ in size). In our data set, cryptobenthic reef fish families included the Gobiidae, Apogonidae, Blenniidae, Tripterygiidae, Bythitidae, and Pseudochromidae, while the large reef fish families included the Pomacentridae, Acanthuridae, Tetraodontidae, Atherinidae, Serranidae, Monacanthidae, Chaetodontidae, Labridae, Muraenidae, and Ophichthidae (Brandl et al. 2018).

Three statistical measures were used to assess alpha diversity: viral abundance (i.e. the standardized number of viral reads), observed viral richness (i.e. the number of viruses), and the Shannon index. We found an association between fish size and observed viral richness, with cryptobenthic reef fishes harbouring more viruses than large reef fishes ($\chi^2 = 2.795$, df = 1, $P = 0.028$) (Fig. 7). Among cryptobenthic reef fishes, *E. tutuilae* contained the most vertebrate-associated viruses with four viral species, followed by *L. lupus* with three viral species (Fig. 6A). To assess whether this difference might be driven by these two outliers, we

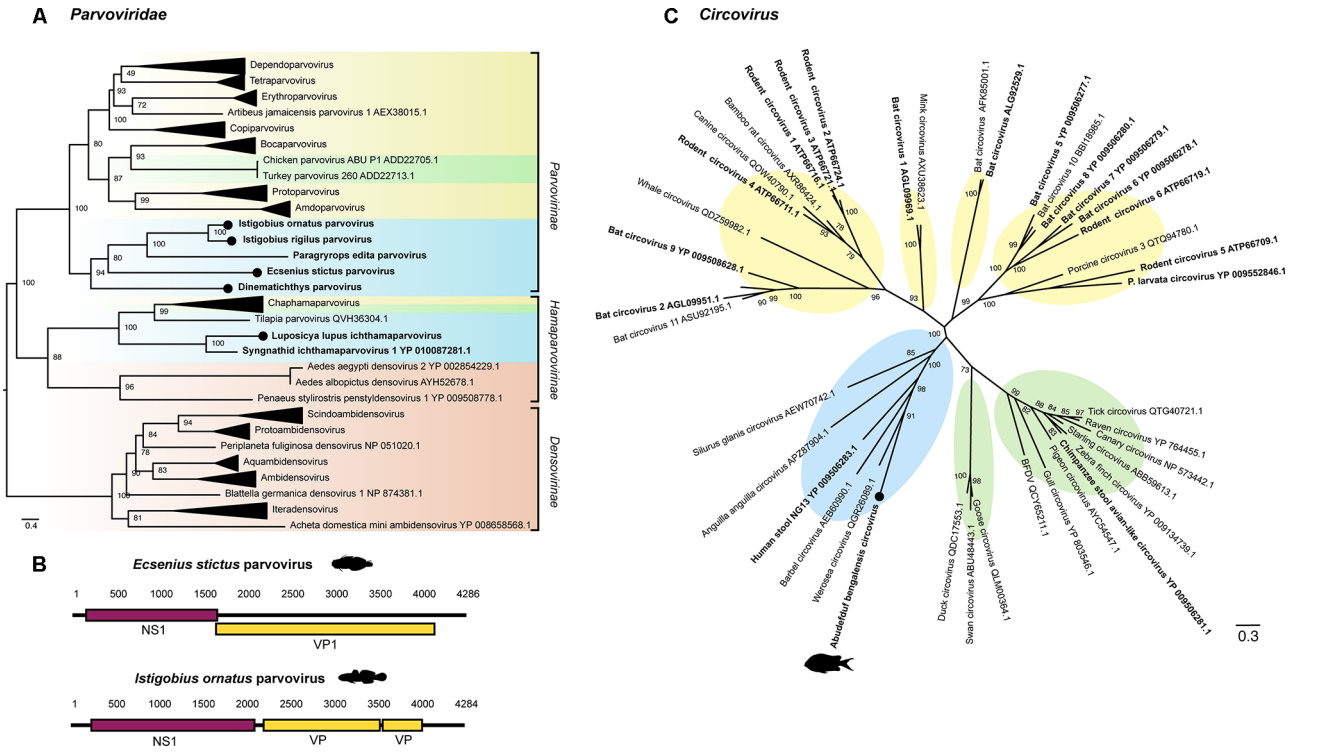

Figure 4. Phylogenetic relationships of the ssDNA viruses identified in this study. (A) The phylogenetic analysis of the NS1 gene of the *Parvoviridae*. Viruses discovered here are represented as black circles. The tree is midpoint rooted for clarity. The scale bar represents the number of amino acid substitutions per site. (B) Schematic genomes (nt) of discovered parvoviruses. Coloured boxes represent both structural (NS1) and non-structural (VP) ORFs. (C) Unrooted phylogeny of the replication-associated protein of the genus *Circovirus* (*Circoviridae*). The circovirus identified is represented as a fish symbol. Branches of both trees are highlighted to represent the host class: blue, fish; red, invertebrates; green, birds; and yellow, mammals. Viruses in the bold text represent sequences identified through metagenomics.

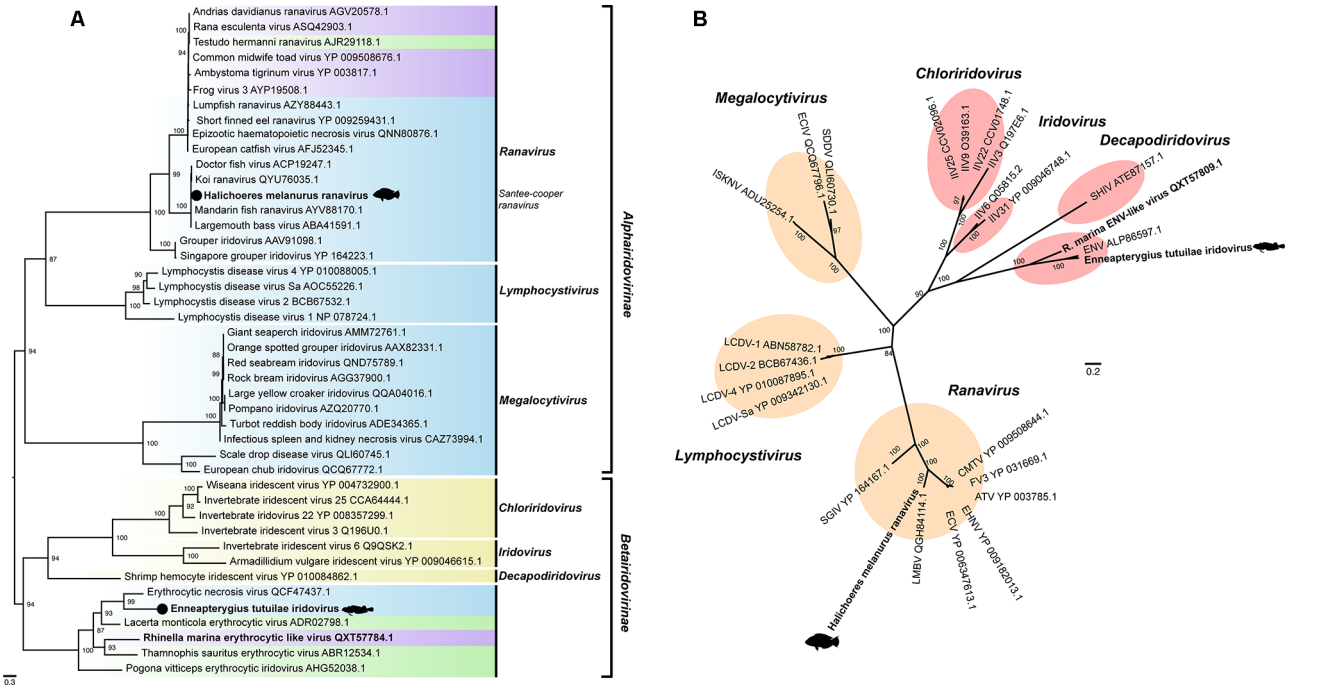

Figure 5. (A) The phylogenetic analysis of the DNA polymerase gene among the *Iridoviridae*. The scale bar represents the number of amino acid substitutions per site. Tip labels represent virus name with NCBI/GenBank accession. Viruses discovered here are represented as a black circle and fish symbol. Branches are highlighted to illustrate host class: blue, fish; purple, amphibians; green, reptiles; and yellow, invertebrates. (B) Unrooted phylogeny of the major capsid protein of the *Iridoviridae*. Branches are coloured to represent both subfamilies: orange, *Alphairidovirinae*; red, *Betairidovirinae*. Viruses in the bold text represent sequences identified through metagenomics.

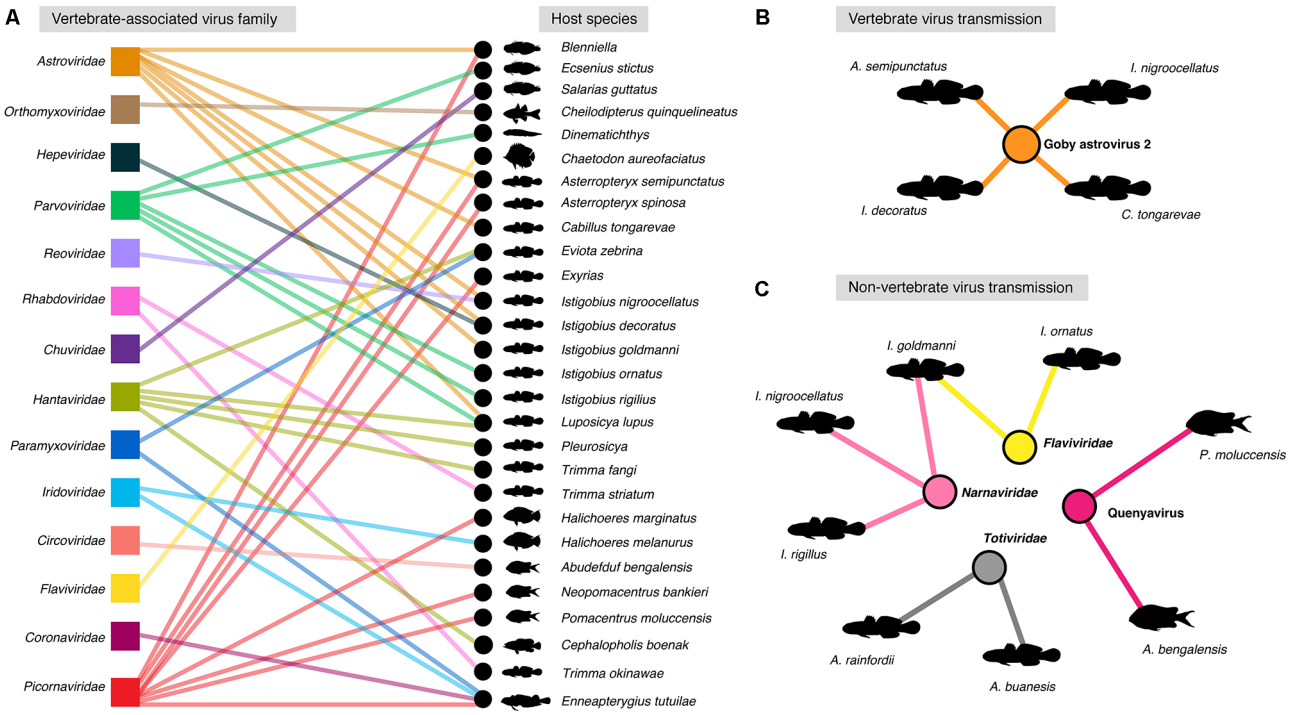

Figure 6. (A) The bipartite network illustrating viral families shared between fish taxa. Viral families are represented as a coloured box, with viruses connected between host species as a coloured line. (B) Network diagram illustrating vertebrate-associated viral species shared between fish. Coloured circle and lines illustrate viral species. (C) Network diagram revealing non-vertebrate viral species shared between fish libraries. All fish silhouettes broadly represent host species at the family level.

omitted these libraries and repeated the analysis. This likewise revealed a significant difference with cryptobenthics exhibiting greater observed viral richness than large reef fish ($\chi^2 = 1.507$, df = 1, $P = 0.044$).

We identified an association between reef fish family and observed viral richness, with the Tripterygiidae exhibiting greater viral diversity than all other reef fish families ($\chi^2 = 16.678$, df = 15, $P = 0.007$). However, we found no association between reef fish family and viral abundance ($P = 0.153$), Shannon diversity ($P = 0.901$), or beta diversity ($R^2 = 0.334$, $P = 0.064$). Likewise, we identified no significant differences in viral abundance ($P = 0.271$), Shannon diversity ($P = 0.142$), nor beta diversity ($R^2 = 0.048$, $P = 0.121$) between cryptobenthic reef fishes and large reef fishes.

As a form of internal control, we repeated our analyses of viral abundance and diversity on the non-vertebrate viruses identified here. This analysis revealed no significant differences in viral abundance between reef fish families ($P = 0.994$) nor between cryptobenthic reef fishes and larger reef fishes ($P = 0.355$). Similarly, we found no significant difference in observed viral richness between fish families ($P = 0.733$) although cryptobenthic reef fishes exhibited higher observed non-vertebrate viral richness than large reef fish families ($\chi^2 = 10.805$, df = 1, $P = 0.016$). We observed no difference in Shannon diversity among fish families ($P = 0.453$), as well as between cryptobenthic reef fishes and larger reef fishes ($P = 0.070$). Similarly, there was no association between beta diversity and reef fish families ($R^2 = 0.279$, $P = 0.126$) nor between cryptobenthic reef fishes and large reef fishes ($R^2 = 0.335$, $P = 0.058$).

We tested for any differences in alpha diversity between dissected and whole fish libraries. This identified no significant differences in vertebrate-associated viral abundance ($P = 0.903$), observed viral richness ($P = 0.058$), or Shannon diversity ($P = 0.156$). Similarly, we detected no significant differences in non-vertebrate viral abundance ($P = 0.055$), observed viral richness ($P = 0.063$), and Shannon diversity ($P = 0.364$). Finally, we tested whether the number of fish per library impacted viral alpha diversity. As expected, this revealed a positive relationship between the number of replicates per fish library and vertebrate viral richness ($R^2 = 0.115$, $P = 0.001$).

Discussion

The GBR supports over 1,200 species of fish and is the largest coral reef ecosystem in the world, comprising 2,500 reefs across approximately 344,400 km^2 (De'ath et al. 2012). Despite the global importance of the GBR, little is known about the natural diversity of viruses that infect tropical reef fishes, as well as the ecological and evolutionary processes that allow such viruses to spread within a reef fish community. To this end, we employed metatranscriptomic sequencing to characterize the viromes of sixty-one tropical reef fish species, including those occupying a 100-m^2 reef fish community and several from the broader surrounding ecosystem. This identified transcripts representing 132 viral sequences, including 38 that exhibited strong phylogenetic relationships with viral genera known to infect vertebrates.

While we sampled coral reef fishes from a small spatial area, we identified a marked absence of cross-species virus transmission, with the only instance of host jumping being the presence of a single viral species (*Astroviridae*) in four goby species (Figs 3C and 6B). However, the observed level of genetic variation (96.8 per cent amino acid similarity) is indicative of past host switching during fish evolution, which may span millions of years, rather than direct host jumping within the ecosystem sampled here.

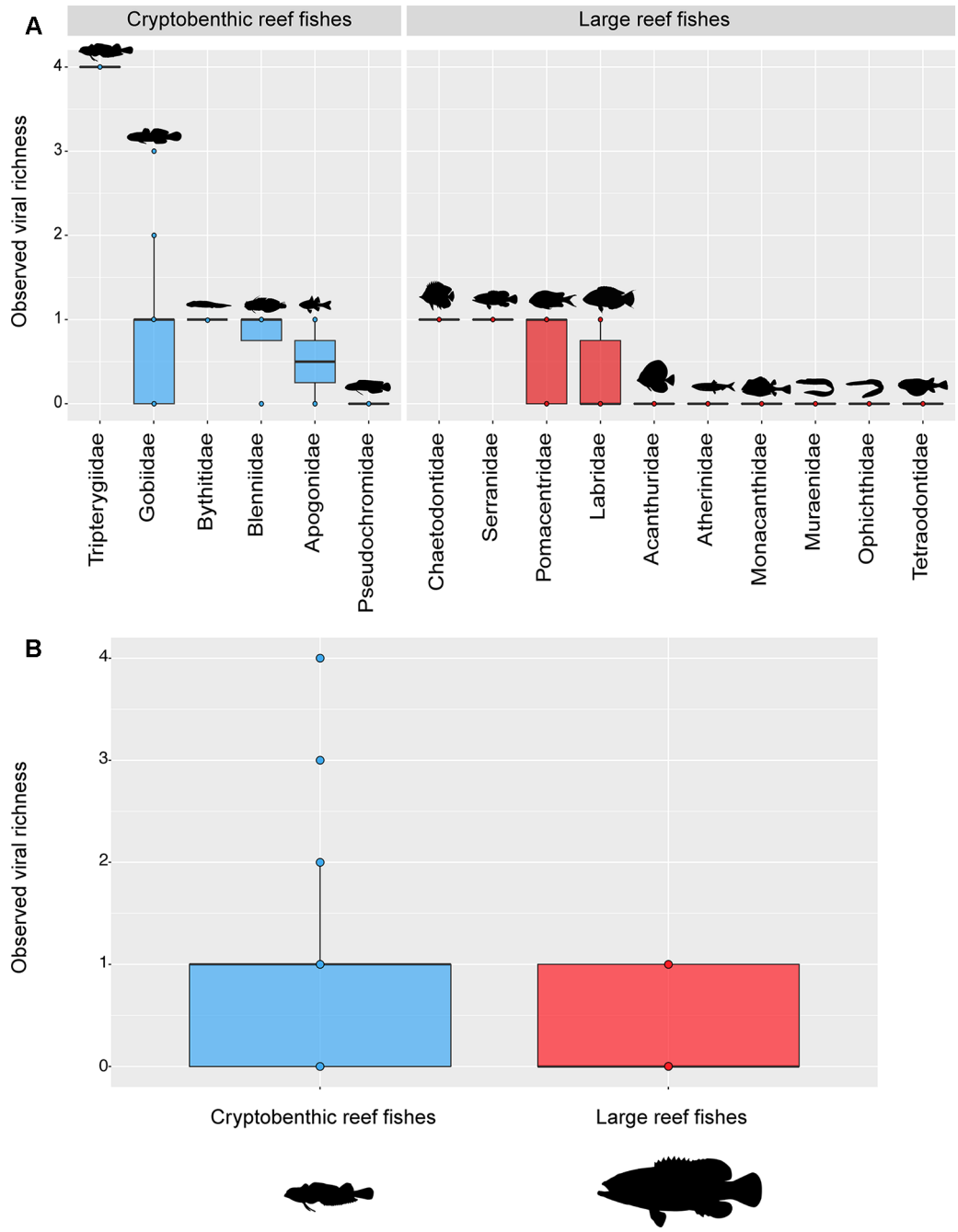

Figure 7. (A) The vertebrate-associated viral richness observed in all tropical reef fish families. (B) The comparison of observed vertebrate-associated viral richness between cryptobenthic reef fishes and large reef fishes. Gobiidae ($n = 30$ species), Labridae ($n = 6$), Pomacentridae ($n = 5$), Blenniidae ($n = 5$), Acanthuridae ($n = 3$), Apogonidae ($n = 2$), Monacanthidae ($n = 2$), Tetraodontidae ($n = 1$), Pseudochromidae ($n = 1$), Chaetodontidae ($n = 1$), Atherinidae ($n = 1$), Serranidae ($n = 1$), Tripterygiidae ($n = 1$), Muraenidae ($n = 1$), Bythitidae ($n = 1$), and Ophichthidae ($n = 1$).

Given the restricted spatial area considered, and the density of potential hosts, such a lack of cross-species transmission suggests that there may be important host barriers (genetic, immune, and physiological) to virus switching among the reef fishes sampled herein. This is supported by the observation that non-vertebrate viruses—that we assume are not impacted by aspects of fish genetics—were characterized by higher levels of cross-species virus transmission. Although more research is needed to understand the nature of these barriers in fish, they may in part reflect differences in host cell receptor binding (Longdon et al. 2014). For instance, although the betacoronavirus RaTG13 sampled from *Rhinolophus affinis* bats is closely related (~96 per cent sequence similarity) to SARS-CoV-2, it is unable to bind to the human ACE2 receptor (Wrobel et al. 2020).

These results also suggest that the reef ecosystem might select for specialist viruses, such that there is no fitness advantage in infecting multiple fish species despite ample exposure. Most fishes within this reef fish community are typically found in high abundances that are sustained year-round (Bellwood et al. 2006; Lefèvre et al. 2016; Brandl et al. 2018). The apparent lack of seasonal variation in fish densities as well as continuously high contact rates and rapid life cycles implies that most of the viruses detected here

have evolved to specialize within a reef fish community and maximize fitness in a single host type. Indeed, specialist-generalist theory suggests that in most host–parasite systems, evolution should select for specialist viruses, particularly when host populations remain stable over space and time (Elena, Agudelo-Romero, and Lalic 2009). Moreover, the evolution of generalist viruses is often restricted by antagonistic pleiotropy, in which mutations that are beneficial in one host species are deleterious in others (Lefeuvre et al. 2019). That we observed very limited evidence for virus generalization across multiple viral families implies that the ecological stability of this reef environment facilitates virus specialization.

We detected differences in vertebrate virome composition between reef fish families, with cryptobenthic reef fishes harbouring more viruses than large reef fishes (Fig. 7). Due to their small body size, cryptobenthic reef fishes exhibit significantly higher rates of metabolism than large reef fishes, resulting in high energy demands with a low tolerance for starvation (Brandl et al. 2018). Given their extremely short lifespans, there are high energy demands for their rapid growth and complex reproductive strategies (e.g. sex change), such that there might be important energy trade-offs between growth, reproduction, and immunity, particularly as immune systems require a substantial allocation of resources for recognizing and eliminating pathogens (Lochmiller and Deerenberg 2000). While it is unclear if and how such energy demands impact immunity in cryptobenthic reef fishes, it is notable that this lifestyle is associated with a more diverse virome than large reef fishes, suggesting that these species may be more susceptible to infection. Indeed, with a daily mortality rate of 8–9 per cent (Depczynski and Bellwood 2005, 2006), investing in immunity may be a poor strategy, and it is possible that these energy-saving strategies are key to the spectacular reproductive success of these cryptobenthic fishes (Brandl et al. 2019). Whether as simple larvae in the pelagic realm, or on the reef, cryptobenthic fishes may be able to minimize energetically expensive activities such as immunity and swimming, freeing up energy to maintain adequate fecundity, despite exceptionally high mortality rates and infection risks.

Among reef fish families, the Tripterygiidae exhibited the greatest observed viral richness, with *E. tutuilae* harbouring four viral species. *Enneapterygius tutuilae* is a generalist on coral reefs, utilizing sand and rubble, soft coral, cave, and open reef microhabitats (Depczynski and Bellwood 2004). It is therefore possible that habitat generalism may increase interactions with other fish species and hence increase the likelihood of being infected by a larger number of viruses although only one example of relatively recent cross-species virus transmission was detected in our system. In contrast, *Gobiodon* species are extreme habitat specialists and may often only occupy a single coral species as habitat for its entire life with minimal interactions with other fish species (Munday, Jones, and Caley 1997). This lifestyle may explain why all five *Gobiodon* species examined here harboured no vertebrate-associated viruses, as well as few non–vertebrate-associated viruses (Fig. 1A–B).

Given the phylogenetic distance between fish and other vertebrate classes, as well as long-term associations through virus–host co-divergence (Zhang et al. 2018), it was not unexpected that 95 per cent of the vertebrate-associated viruses discovered here clustered with other fish viruses. A case in point comes from the phylogeny of the *Hantaviridae* that clearly reflects the broad evolutionary history of vertebrates, with the *Agantavirinae* (jawless fish) falling basal to the *Actantavirinae* (ray-finned fish), *Reptantavirinae* (reptiles), and *Mammantavirinae* (mammals), respectively (Fig. 2). Despite their likely long-term presence in ray-finned fish, actinoviruses have only recently been discovered and have been associated with disease in farmed species (Shi et al. 2018; Geoghegan et al. 2021; Hierweger et al. 2021). For instance, a novel actinovirus—perch actinovirus—was identified in diseased European perch (*Perca fluviatilis*) with high concentrations of viral RNA in gill endothelial cells and macrophages (Hierweger et al. 2021). Our discovery of four actinoviruses in gobies suggests that these viruses may be widespread in this family and hence should be monitored closely if interacting with farmed populations.

Another virus of concern identified in our study is a novel isolate of *Santee-Cooper ranavirus* in *H. melanurus*. To the best of our knowledge, this is the first discovery of this virus in Australia. Our detection of this virus in seemingly healthy wrasses, as well as its natural presence in cleaner wrasses (*Labroides dimidiatus*), suggests that this family of tropical reef fishes may be important reservoir hosts for the *Santee-Cooper ranavirus* group (Hedricks and McDowell 1995). For instance, disease outbreaks with high mortality rates have only been observed in farmed freshwater fish species including largemouth bass (*Micropterus salmoides*), mandarinfish (*Siniperca chuatsi*), and koi (*Cyprinus carpio*) (Plumb et al. 1996; Grizzle et al. 2002; George et al. 2015; Zhang et al. 2020; Zhao et al. 2020).

Despite the long-known presence of viral erythrocytic necrosis in various fish species across the North Atlantic and Pacific Oceans, the genome of the causative virus—ENV—was only recently characterized (Pagowski et al. 2019). We identified a closely related virus in *E. tutuilae*. Viral erythrocytic necrosis in a juvenile triggerfish (*Rhinecanthus aculeatus*) has been reported nearby at Lizard Island (Davies et al. 2009), suggesting that this virus may be circulating between several reef locations along the GBR. Phylogenetic comparisons of DNA polymerase and major capsid protein revealed a distinct 'ectothermic vertebrate' clade that could be clearly classified as a novel genus within the *Betairidovirinae, Iridoviridae* (Fig. 5).

While we observed significant differences in virome composition between coral reef fishes, there were necessary limitations in our sampling. For instance, cryptobenthic fishes made up 62 per cent of the fish diversity in this study. In addition, twenty-three libraries contained only one individual. Such unequal sample sizes likely impacted the statistical power of our analyses. Accordingly, where possible, future work should balance the number of samples as well as increase the number of reef sampling sites. Another limitation was the difficulty in detecting highly divergent viral sequences with the similarity-based methods used here. Moreover, given the RNA-based nature of metatranscriptomics, we were only able to detect DNA viruses that were actively generating transcripts, in turn reducing the chances of identifying viruses that may be expressing genes at lower frequencies.

In sum, our study identified a large diversity of viruses in tropical reef fish assemblages. We identified a marked absence of virus exchange within a reef fish community, suggesting that there may be important host barriers for successful cross-species virus transmission. Our discovery of a novel *Santee-Cooper ranavirus* isolate in seemingly healthy wrasses highlights the importance of virological surveillance in marine wildlife, particularly as this virus can cause significantly high mortality in farmed fishes. As such, these species should be considered in biosecurity risk assessments and screened if utilized for aquaculture or aquarium operations (Rimmer et al. 2015). Accordingly, future studies should also investigate its susceptibility in Australian food fish to fully assess its threat of emergence (Hedricks and McDowell 1995). Overall, this study increases our knowledge on the severely

understudied coral reef fish virome and provides the first data on virus–host interactions in a reef fish community.

Data availability

All sequence reads are available on the NCBI Sequence Read Archive under BioProject PRJNA841039, and all generated virus sequences have been deposited in NCBI/GenBank under accessions ON595936–ON596019.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Funding

This work was funded by the Australian Research Council (ARC) Australian Laureate Fellowships to E.C.H. (FL170100022) and D.R.B. (FL190100062) and an ARC Discovery Project grant to E.C.H. and J.L.G. (DP200102351). J.L.G. is funded by a New Zealand Royal Society Rutherford Discovery Fellowship (RDF-20-UOO-007) and a Marsden Fast Start grant (20-UOO-105).

Conflict of interest: The authors declare no competing interests.

References

Ackerman, J. L., and Bellwood, D. R. (2002) 'Comparative Efficiency of Clove Oil and Rotenone for Sampling Tropical Reef Fish Assemblages', *Journal of Fish Biology*, 60: 893–901.

Allison, A. B. et al. (2011) 'Characterization of Durham Virus, a Novel Rhabdovirus that Encodes Both a C and SH Protein', *Virus Research*, 155: 112–22.

Alves de Matos, A. P. et al. (2010) 'New Viruses from *Lacerta monticola* (Serra da Estrela, Portugal): Further Evidence for a New Group of Nucleo-Cytoplasmic Large Deoxyriboviruses', *Microscopy and Microanalysis*, 17: 101–8.

Barton, J. A. et al. (2020) 'Biological Controls to Manage Acropora-eating Flatworms in Coral Aquaculture', *Aquaculture Environment Interactions*, 12: 61–6.

Bellwood, D. R. et al. (2004) 'Confronting the Coral Reef Crisis', *Nature*, 429: 827–33.

Bellwood, D. R. et al. (2006) 'Coral Bleaching, Reef Fish Community Phase Shifts and the Resilience of Coral Reefs', *Global Change Biology*, 12: 1587–94.

Bolger, A. M., Lohse, M., and Usadel, B. (2014) 'Trimmomatic: A Flexible Trimmer for Illumina Sequence Data', *Bioinformatics*, 30: 2114–20.

Booth, D. J. et al. (2018) 'Predicting Success of Range-expanding Coral Reef Fish in Temperate Habitats Using Temperature-abundance Relationships', *Frontiers in Marine Science*, 5: 31.

Brandl, S. J. et al. (2018) 'The Hidden Half: Ecology and Evolution of Cryptobenthic Fishes on Coral Reefs', *Biological Reviews of the Cambridge Philosophical Society*, 93: 1846–73.

Brandl, S. J. et al. (2019) 'Demographic Dynamics of the Smallest Marine Vertebrates Fuel Coral Reef Ecosystem Functioning', *Science*, 364: 1189–92.

Buchfink, B., Xie, C., and Huson, D. H. (2015) 'Fast and Sensitive Protein Alignment Using DIAMOND', *Nature Methods*, 12: 59–60.

Burge, C. A. et al. (2014) 'Climate Change Influences on Marine Infectious Diseases: Implications for Management and Society', *Annual Review of Marine Science*, 6: 249–77.

Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. (2009) 'trimAl: A Tool for Automated Alignment Trimming in Large-scale Phylogenetic Analyses', *Bioinformatics*, 25: 1972–3.

Costa, V. A. et al. (2021) 'Metagenomic Sequencing Reveals a Lack of Virus Exchange between Native and Invasive Freshwater Fish across the Murray–Darling Basin, Australia', *Virus Evolution*, 7: veab034.

Davies, A. J. et al. (2009) 'Suspected Viral Erythrocytic Necrosis (VEN) in a Juvenile Blackbar Triggerfish, *Rhinecanthus aculeatus*, from Lizard Island, Great Barrier Reef, Australia', *Marine Biodiversity Records*, 2: E149.

De'ath, G. et al. (2012) 'The 27-year Decline of Coral Cover on the Great Barrier Reef and Its Causes', *Proceedings of the National Academy of Sciences*, 109: 17995–9.

Depczynski, M., and Bellwood, D. R. (2003) 'The Role of Cryptobenthic Reef Fishes in Coral Reef Trophodynamics', *Marine Ecology Progress Series*, 256: 183–91.

——— (2004) 'Microhabitat Utilisation Patterns in Cryptobenthic Coral Reef Fish Communities', *Marine Biology*, 145: 455–63.

——— (2005) 'Shortest Recorded Vertebrate Lifespan Found in a Coral Reef Fish', *Current Biology*, 15: 288–9.

——— (2006) 'Extremes, Plasticity, and Invariance in Vertebrate Life History Traits: Insights from Coral Reef Fishes', *Ecology*, 87: 3119–27.

Dixon, P. (2003) 'VEGAN, a Package of R Functions for Community Ecology', *Journal of Vegetation Science*, 14: 927–30.

Elena, S. F., Agudelo-Romero, P., and Lalic, J. (2009) 'The Evolution of Viruses in Multi-host Fitness Landscapes', *The Open Virology Journal*, 3: 1–6.

FAO. (2020), *The State of World Fisheries and Aquaculture. Sustainability in Action* <https://www.fao.org/documents/card/en/c/ca9229en/> accessed 9 Feb 2022.

Geoghegan, J. L. et al. (2018) 'Hidden diversity and evolution of viruses in market fish', *Virus Evolution*, 4: vey031.

——— (2021) 'Virome Composition in Marine Fish Revealed by Meta-transcriptomics', *Virus Evolution*, 7: veab005.

George, M. R. et al. (2015) 'Isolation and Characterization of a Ranavirus from Koi, *Cyprinus carpio* L., Experiencing Mass Mortalities in India', *Journal of Fish Diseases*, 38: 389–403.

Go, J. et al. (2006) 'The Molecular Epidemiology of Iridovirus in Murray Cod (*Maccullochella peelii peelii*) 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.

Grizzle, J. M. et al. (2002) 'First Isolation of Largemouth Bass Virus', *Diseases of Aquatic Organisms*, 50: 233–5.

Grosset, C. et al. (2014) 'Intraerythrocytic Iridovirus in Central Bearded Dragons (*Pogona vitticeps*)', *Journal of Veterinary Diagnostic Investigation*, 26: 354–64.

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: 1494–512.

Hedricks, R. P., and McDowell, T. S. (1995) 'Properties of Iridoviruses from Ornamental Fish', *Veterinary Research*, 26: 423–7.

Heron, S. F. et al. (2016) 'Warming Trends and Bleaching Stress of the World's Coral Reefs 1985-2012', *Scientific Reports*, 6: 38402.

Hierweger, M. M. et al. (2021) 'Novel Filoviruses, Hantavirus, and Rhabdovirus in Freshwater Fish, Switzerland, 2017', *Emerging Infectious Diseases*, 27: 3082–91.

Hothorn, T., Bretz, F., and Westfall, P. (2008) 'Simultaneous Inference in General Parametric Models', *Biometrical Journal*, 50: 346–63.

Johnson, W. E. (2019) 'Origins and Evolutionary Consequences of Ancient Endogenous Retroviruses', *Nature Reviews Microbiology*, 17: 355–70.

Jones, P. et al. (2014) 'InterProScan 5: Genome-scale Protein Function Classification', *Bioinformatics*, 30: 1236–40.

Kalyaanamoorthy, 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.

Kibenge, F. S. B. (2019) 'Emerging Viruses in Aquaculture', *Current Opinion in Virology*, 34: 97–103.

Kurz, W. et al. (1986) 'Isolation and Characterization of a Tupaia Rhabdovirus', *Intervirology*, 25: 88–96.

Lagkouvardos, I. et al. (2017) 'Rhea: A Transparent and Modular R Pipeline for Microbial Profiling Based on 16S rRNA Gene Amplicons', *PeerJ*, 5: e2836.

Lam, V. W. Y. et al. (2020) 'Climate Change, Tropical Fisheries and Prospects for Sustainable Development', *Nature Reviews Earth & Environment*, 1: 440–54.

Lancaster, M. J., Williamson, M. M., and Schroen, C. J. (2003) 'Iridovirus-associated Mortality in Farmed Murray Cod (*Maccullochella peelii peelii*)', *Australian Veterinary Journal*, 81: 633–4.

Langmead, B., and Salzberg, S. L. (2012) 'Fast Gapped-read Alignment with Bowtie 2', *Nature Methods*, 9: 357–9.

Lefeuvre, P. et al. (2019) 'Evolution and Ecology of Plant Viruses', *Nature Reviews Microbiology*, 17: 632–44.

Lefèvre, C. D. et al. (2016) 'Consequences of Extreme Life History Traits on Population Persistence: Do Short-lived Gobies Face Demographic Bottlenecks?', *Coral Reefs*, 35: 399–409.

Li, C. X. et al. (2015a) 'Unprecedented Genomic Diversity of RNA Viruses in Arthropods Reveals the Ancestry of Negative-sense RNA Viruses', *eLife*, 4: e05378.

Li, D. et al. (2015b) 'MEGAHIT: An Ultra-fast Single-node Solution for Large and Complex Metagenomics Assembly via Succinct de Bruijn Graph', *Bioinformatics*, 31: 1674–6.

Lochmiller, R. L., and Deerenberg, C. (2000) 'Trade-offs in Evolutionary Immunology: Just What Is the Cost of Immunity?', *Oikos*, 88: 87–98.

Longdon, B. et al. (2014) 'The Evolution and Genetics of Virus Host Shifts', *PLoS Pathogens*, 10: e1004395.

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.

Moberg, F., and Folke, C. (1999) 'Ecological Goods and Services of Coral Reef Ecosystems', *Ecological Economics*, 29: 215–33.

Munday, P. L., Jones, G. P., and Caley, M. J. (1997) 'Habitat Specialisation and the Distribution and Abundance of Coral-dwelling Gobies', *Marine Ecology Progress Series*, 152: 227–39.

Murray, A. G. (2016) 'A Modelling Framework for Assessing the Risk of Emerging Diseases Associated with the Use of Cleaner Fish to Control Parasitic Sea Lice on Salmon Farms', *Transboundary and Emerging Diseases*, 63: 270–7.

Nayfach, S. et al. (2021) 'CheckV Assesses the Quality and Completeness of Metagenome-assembled Viral Genomes', *Nature Biotechnology*, 39: 578–85.

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.

Nunes, M. R. T. et al. (2017) 'Genetic Characterization, Molecular Epidemiology, and Phylogenetic Relationships of Insect-specific Viruses in the Taxon Negevirus', *Virology*, 504: 152–67.

Obbard, D. J. et al. (2020) 'A New Lineage of Segmented RNA Viruses Infecting Animals', *Virus Evolution*, 6: vez061.

Ostfeld, R. S., and Keesing, F. (2012) 'Effects of Host Diversity on Infectious Disease', *Annual Review of Ecology, Evolution, and Systematics*, 43: 157–82.

Pagowski, V. A. et al. (2019) 'Distribution and Phylogeny of Erythrocytic Necrosis Virus (ENV) in Salmon Suggests Marine Origin', *Viruses*, 11: 358.

Panek, F. M. (2005) 'Epizootics and Disease of Coral Reef Fish in the Tropical Western Atlantic and Gulf of Mexico', *Reviews in Fisheries Science*, 13: 1–21.

Parrish, C. R. et al. (2008) 'Cross-species Virus Transmission and the Emergence of New Epidemic Diseases', *Microbiology and Molecular Biology Reviews*, 72: 457–70.

Pénzes, J. J. et al. (2020) 'Reorganizing the Family *Parvoviridae*: A Revised Taxonomy Independent of the Canonical Approach Based on Host Association', *Archives of Virology*, 165: 2133–46.

Plumb, J. A. et al. (1996) 'An Iridovirus Isolated from Wild Largemouth Bass', *Journal of Aquatic Animal Health*, 8: 265–70.

Reuter, G. et al. (2015) 'A Highly Divergent Picornavirus in an Amphibian, the Smooth Newt (*Lissotriton vulgaris*)', *Journal of General Virology*, 96: 2607–13.

Rimmer, A. E. et al. (2015) 'Detection of Dwarf Gourami Iridovirus (*Infectious Spleen and Kidney Necrosis Virus*) in Populations of Ornamental Fish Prior to and after Importation into Australia, with the First Evidence of Infection in Domestically Farmed Platy (*Xiphophorus maculatus*)', *Preventive Veterinary Medicine*, 122: 181–94.

Russo, A. G. et al. (2021) 'Discovery of Novel Viruses Associated with the Invasive Cane Toad (*Rhinella marina*) in Its Native and Introduced Ranges', *Frontiers in Microbiology*, 12: 733631.

Santini, L. et al. (2018) 'Global Drivers of Population Density in Terrestrial Vertebrates', *Global Ecology and Biogeography*, 27: 968–79.

Schmidt, K. A., and Ostfeld, R. S. (2001) 'Biodiversity and the Dilution Effect in Disease Ecology', *Ecology*, 82: 609–19.

Shi, M. et al. (2016) 'Redefining the Invertebrate RNA Virosphere', *Nature*, 540: 539–43.

Shi, M. et al. (2018) 'The Evolutionary History of Vertebrate RNA Viruses', *Nature*, 556: 197–202.

Spalding, M. D., and Grenfell, A. M. (1997) 'New Estimates of Global and Regional Coral Reef Areas', *Coral Reefs*, 16: 225–30.

Stuart-Smith, R. D. et al. (2018) 'Ecosystem Restructuring along the Great Barrier Reef following Mass Coral Bleaching', *Nature*, 560: 92–6.

Thurber, R. V. et al. (2017) 'Virus-host Interactions and Their Roles in Coral Reef Health and Disease', *Nature Reviews Microbiology*, 15: 205–16.

Untergasser, A. et al. (2012) 'Primer3—New Capabilities and Interfaces', *Nucleic Acids Research*, 40: e115.

Valero-Mora, P. M. (2010) 'ggplot2:ElegantGraphics for Data Analysis', *Journal of Statistical Software*, 35: 1–3.

Vergés, A. et al. (2014) 'The Tropicalization of Temperate Marine Ecosystems: Climate-mediated Changes in Herbivory and Community Phase Shifts', *Proceedings of the Royal Society B*, 281: 20140846.

Wellehan, J. F. X. et al. (2008) 'Characterization of an Erythrocytic Virus in the Family Iridoviridae from a Peninsula Ribbon Snake (*Thamnophis sauritus sackenii*)', *Veterinary Microbiology*, 131: 115–22.

Whittington, R. J., Becker, J. A., and Dennis, M. M. (2010) 'Iridovirus Infections in Finfish—Critical Review with Emphasis on Ranaviruses', *Journal of Fish Diseases*, 33: 95–122.

Woodhead, A. J. et al. (2019) 'Coral Reef Ecosystem Services in the Anthropocene', *Functional Ecology*, 33: 1023–34.

Wrobel, A. G. et al. (2020) 'SARS-CoV-2 and Bat RaTG13 Spike Glycoprotein Structures Inform on Virus Evolution and Furin-cleavage Effects', *Nature Structural & Molecular Biology*, 27: 763–7.

Zhang, Y. Z. et al. (2018) 'The Diversity, Evolution and Origins of Vertebrate RNA Viruses', *Current Opinion in Virology*, 31: 9–16.

Zhang, W. et al. (2020) 'Widespread Outbreaks of the Emerging Mandarinfish Ranavirus (MRV) Both in Natural and ISKNV-FKC Vaccinated Mandarinfish *Siniperca chuatsi* in Guangdong, South China, 2017', *Aquaculture*, 520: 734989.

Zhao, R. et al. (2020) 'A New Ranavirus of the Santee-Cooper Group Invades Largemouth Bass (*Micropterus salmoides*) Culture in Southwest China', *Aquaculture*, 526: 735363.