An Unconventional Flavivirus and other RNA Viruses in the Sea Cucumber (Holothuroidea; Echinodermata) Virome

Ian Hewson1*, Mitchell R. Johnson2, Ian R. Tibbetts3

1 Department of Microbiology, Cornell University; hewson@cornell.edu
2 Department of Microbiology, Cornell University; mrj78@cornell.edu
3 School of Biological Sciences, University of Queensland; i.tibbetts@uq.edu.au

* Correspondence: hewson@cornell.edu; Tel.: +1-607-255-0151

Abstract: Sea cucumbers (Holothuroidea; Echinodermata) are ecologically significant constituents of benthic marine habitats. We surveilled RNA viruses inhabiting 8 species (representing 4 families) of holothurian collected from four geographically distinct locations by viral metagenomics, including a single specimen of Apostichopus californicus affected by a hitherto undocumented wasting disease. The RNA virome comprised genome fragments of both single-stranded positive sense and double stranded RNA viruses, including those assigned to the Picornavirales, Ghabrivirales, and Amarillovirales. We discovered an unconventional flavivirus genome fragment which was most similar to a shark virus. Ghabivirales-like genome fragments were most similar to fungal totiviruses in both genome architecture and homology, and likely infected mycobiome constituents. Picornavirales, which are commonly retrieved in host-associated viral metagenomes, were similar to invertebrate transcriptome-derived picorna-like viruses. Sequence reads recruited from the grossly normal A. californicus metavirome to nearly all viral genome fragments recovered from the wasting-affected A. californicus. The greatest number of viral genome fragments was recovered from wasting A. californicus compared to any other surveyed holothurian, including the grossly normal A. californicus, which reflects a pattern observed in viral metagenomics study of sea star wasting. These results expand the known host range of flaviviruses, and suggest that fungi and their viruses may play roles in holothurian ecology.

Keywords: holothurian; Apostichopus; wasting; virus; flavivirus; totivirus

1. Introduction

Next generation DNA sequencing technology applied to viral metagenomics has enabled surveillance of viruses associated with invertebrate tissues. These studies, along with mining metazoan transcriptomes, led to the discovery of novel viral lineages in aquatic invertebrates and broadened the host range of several viral families [1-15]. While the majority of viral surveillance and discovery is focused on grossly normal individual specimens of aquatic metazoa, viral metagenomics has been used to examine the presence of potential pathogenic viruses through comparative asymptomatic/disease affected individuals (reviewed in [16]). Despite a growing appreciation for the diversity of aquatic metazoan-associated viruses, and of their potential roles in host disease, they remain largely undersampled. Study of novel and/or highly divergent viral genomes across a wider range of aquatic invertebrates may provide clues to viral evolution and potential roles in host ecology.

Holothurians (Holothuroidea; Echinodermata) are ecologically important echinoderms [17, 18] that can dominate benthic biomass, and contribute significantly to biogeochemical cycling in benthic compartments. On the abyssal plains, for example, holothurian biomass exceeds all other invertebrates [19]. Holothurians may be planktivorous, herbivorous or deposit feeders, and
contribute to ecosystem function by regeneration of particulate organic matter into dissolved organic and inorganic constituents [18, 20]. Feeding by holothurians on sediments stimulates bacterial production [20] and potentially bacterial diversity [21]. Holothurians produce pelagic larvae that as meroplankton contribute to herbivory [22-24]. Holothurians are also economically significant, since they are fished and aquacultured for human consumption [25]. In China, the world’s largest consumer of holothurians, more than 200 kT worth CNY 176M (US$26M) were produced in 2017 [26]. The economic, evolutionary and ecological significance of holothurians make these groups attractive to study of factors influencing their biology and population dynamics, including potentially pathogenic microorganisms.

Like all echinoderms, holothurians are deuterostomes and are more therefore more closely related to chordates than other invertebrate groups. From a viral perspective, their similarity to chordates, and specifically their similarity of cell surface properties that mediate endocytosis and fusion, suggest they may be infected by similar viral groups. For example, sialic acids, which are used by a variety of viral families to enter cells [27], became prominent late in evolution, especially in deuterostomes, where they play diverse physiological roles [28]. While echinoderms lack adaptive immunity, genome analyses of the echinoid Strongylocentrotus purpuratus revealed homologs of vertebrate immune factors [29, 30]. Despite these similarities, few studies have assessed the composition of viruses in echinoderm tissues. DNA viruses, including Piccovirales [11, 12, 14] and Curlivirales [31, 32] were discovered by viral metagenomic studies using amplified (φ29-mediated rolling circle replication) material. RNA viral surveys, which demand a different amplification protocol and handling from DNA viruses, have detected Baphyvirales, Picornavirales, Articulavirales and Tolivirales [13, 33] in asteroids. However, RNA viruses in other echinoderm classes have not been investigated by viral metagenomic approaches.

In this study we surveilled RNA viruses in eight holothurian species collected from four distinct geographic locations (Heron Island, Moreton Bay, Salish Sea and Southeast Alaska) and representing four families of holothurians. We found that viruses associated with these tissues were dominated by Picornavirales (Marnaviridae) and Ghabrivirales (Totiviridae), but report on the presence on a deeply branched flavivirus that was present in two holothurian species in the northeast Pacific Ocean. We also compared the composition of viruses associated with a grossly normal Apostichopus californicus and a specimen that was affected by sea cucumber wasting (a condition that is not extensively documented in the peer-reviewed literature), but found little difference in viral genome representation.

2. Materials and Methods

Specimen Collection: Holothurian specimens were collected at four geographic locations between 2015 and 2017 (Table 1). Specimens were collected by hand in the intertidal zone or by SCUBA Diver (with the exception of Cucumaria miniata, which was obtained by rock dredge). Specimens of Apostichopus californicus (renamed from Parastichopus californicus; [34]) were collected by commercial fishers and retrieved by the Alaska Department of Fish and Game. The presence of a wasting-like condition was determined by gross disease signs on collection (see below). Specimens were subsampled either immediately after collection, or whole specimens frozen on dry ice or liquid N₂ before transport to the laboratory at Cornell University, where they were thawed and dissected prior to metavirome preparation.
Table 1. Specimens collected for viral metagenomics survey and library characteristics. The * indicates the specimen was affected by sea cucumber wasting (SCW).

<table>
<thead>
<tr>
<th>Species</th>
<th>Family</th>
<th>Sample Location</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Collection Date</th>
<th>Sequence Library Size (reads)</th>
<th># Contigs</th>
<th>N50 contigs</th>
<th>Contigs Matching Viral Genomes E&lt;10^{-20}</th>
<th>Reads Mapped to Viral Contigs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cucumaria miniata</td>
<td>Cucumariidae</td>
<td>Salish Sea</td>
<td>48.233 N</td>
<td>122.7917 W</td>
<td>7-Jan-16</td>
<td>1686849</td>
<td>5394</td>
<td>811</td>
<td>3</td>
<td>247494</td>
</tr>
<tr>
<td>Apostichopus californicus</td>
<td>Stichopodidae</td>
<td>Ketchikan</td>
<td>55.4351 N</td>
<td>131.9481 W</td>
<td>26-Oct-16</td>
<td>897618</td>
<td>3431</td>
<td>804</td>
<td>11</td>
<td>195230</td>
</tr>
<tr>
<td>Apostichopus californicus*</td>
<td>Stichopodidae</td>
<td>Ketchikan</td>
<td>55.4351 N</td>
<td>131.9481 W</td>
<td>26-Oct-16</td>
<td>1019071</td>
<td>3900</td>
<td>880</td>
<td>2</td>
<td>179175</td>
</tr>
<tr>
<td>Holothuria scabra</td>
<td>Holothuridae</td>
<td>Amity Banks</td>
<td>27.4115 S</td>
<td>153.4344 E</td>
<td>10-Dec-15</td>
<td>822244</td>
<td>21777</td>
<td>838</td>
<td>1</td>
<td>368</td>
</tr>
<tr>
<td>Synaptula recta</td>
<td>Synaptidae</td>
<td>Dunwich</td>
<td>27.4952 S</td>
<td>153.3985 E</td>
<td>10-Dec-15</td>
<td>983151</td>
<td>19760</td>
<td>847</td>
<td>0</td>
<td>299</td>
</tr>
<tr>
<td>Holothuria diffcilis</td>
<td>Holothuridae</td>
<td>Amity Banks</td>
<td>27.4115 S</td>
<td>153.4344 E</td>
<td>10-Dec-15</td>
<td>794579</td>
<td>5803</td>
<td>846</td>
<td>1</td>
<td>301</td>
</tr>
<tr>
<td>Stichopus horrens</td>
<td>Stichopodidae</td>
<td>Amity Banks</td>
<td>27.4115 S</td>
<td>153.4346 E</td>
<td>10-Dec-15</td>
<td>1577532</td>
<td>21194</td>
<td>817</td>
<td>2</td>
<td>423</td>
</tr>
<tr>
<td>Holothuria atra</td>
<td>Holothuridae</td>
<td>Heron Island</td>
<td>23.4441 S</td>
<td>151.9113 E</td>
<td>19-Dec-15</td>
<td>875840</td>
<td>18866</td>
<td>793</td>
<td>1</td>
<td>345</td>
</tr>
<tr>
<td>Holothuria pardalis</td>
<td>Holothuridae</td>
<td>Dunwich</td>
<td>27.4952 S</td>
<td>153.3985 E</td>
<td>10-Dec-15</td>
<td>1467795</td>
<td>7424</td>
<td>896</td>
<td>0</td>
<td>306</td>
</tr>
</tbody>
</table>
Metavirome preparation: Metaviromes were prepared from each specimen by first retrieving an 8 mm biopsy punch from body wall tissues, which was extruded into a sterile microcentrifuge tube. From here, viral metagenomes were prepared according to established protocols as described in [35] with modifications [36]. Briefly, tissue samples were disrupted by bead beating (Zymo Bead Beaters) in 2 mL virus-free PBS, then centrifuged at 5,000 x g for 5 min to remove large particulate material. The resulting supernatant was then removed, filtered through 0.2 µm PES filters (to remove host cells and cell debris), and then treated with RNaseOne (50 U), DNase I (5 U), and Benzonase nuclease (250 U) for 2 h at 37°C. Nuclease activity was arrested by amendment with 50 µM EDTA. RNA was extracted from purified virus-sized material using the Zymo Viral RNA Kit. Extracted RNA was then amplified using the TransPlex WTA2 kit (Sigma Aldrich) applied to 5 µL extract. Because we did not standardize template amounts, and because of uncertainties in bias introduced by TransPlex amplification, data presented in this manuscript are not quantitative and based solely on presence and absence of genome fragments [16]. Amplified DNA (the end product of the TransPlex protocol) was submitted to the Cornell University Biotechnology Resource Center, where it was prepared with the Nextera XT library preparation protocol, and sequenced on an Illumina MiSeq 2 x 250 bp platform. Sequence libraries are available at NCBI under accession PRJNA417963.

Viral Metagenome Library Analyses: Sequence libraries were first trimmed for ambiguous bases, adapters, TransPlex primers, and poor quality (N>2) sequences using the Trim Sequences function in the CLC Genomics Workbench 4.0. Sequence libraries were then assembled using a minimum overlap of 0.2 and similarity of 0.9 resulting in contig spectra for each library. The resulting contig spectra was aligned against several boutique databases of RNA viruses: 1) all RNA viral library using tBLASTx (genome sequences retrieved in September 2018 from NCBI using search term “RNA Virus”); all Mononegavirus proteins (search term “mononegaviruses”) by BLASTx; Picornavirus ribosome-dependent RNA polymerase (RdRp) proteins (search term “RdRp AND picornavirus”) by BLASTx; invertebrate RNA viral proteins (search term “invertebrate AND RNA viruses”) by BLASTx; Flavivirus proteins (search term “flavivirus”) by BLASTx; Coronavirus proteins (search term “coronavirus”) by BLASTx; and nodavirus proteins (search term “nodavirus”) by BLASTx. Sequences were further vetted by comparing matches to viral proteins against the non-redundant database at NCBI by BLASTn, and matches (E<10^-20) to bacterial or eukaryotic genes were removed from further consideration. Finally, contigs were compared to the RefSeq and non-redundant protein database by BLASTx to retrieve closest matches from cultivated and uncultivated viruses. Viral contiguous sequences are available at NCBI under accessions MT949664 - MT949685.

Viral Genome Architecture and Phylogeny: The open reading frames on vetted viral contigs were extracted with GetORF [37] and compared against established genome architecture of viral families. If significant matches against proteins were detected, but ORF lengths were shorter than expected based on reported genome architectures, sequence was examined for internal frameshifts, which are common in viral ORFs. ORFs were then annotated by comparison against the RefSeq Viral Proteins Database (NCBI). ORFs and partial ORFs adjacent to annotated viral proteins were further compared against the non-redundant database at NCBI by tBLASTx [38], and if no matches were found, then by the Phyre protein server [39]. Internal secondary RNA structures corresponding with published viral genomes were checked using the mFold server [40]. Additional features were identified by the conserved domain database (CDD) at NCBI [41]. α-helix and β-strand regions of ORFs were determined by PSIPRED [42]. Viral contig sequences and closest matches in NCBI were collected and aligned using the native alignment algorithm in CLC Sequence Viewer 8.0 (gap open cost 10.0; gap extension cost 1; end gap cost as any other), and trimmed to contiguous sequences within protein encoding regions.

Eukaryotic 18S and 28S rRNA Analyses: 18S and 28S rRNAs were identified in viral metagenomes by comparison of contig spectra against the Silva database [43] by BLASTn (E<10^-20). Contigs matching this criteria were further annotated based on nearest match in NCBI by BLASTn, and classified based on taxonomic descriptions provided by the NCBI Taxonomy Database.

Sea cucumber wasting disease: A condition affecting A. californicus has been observed since at least 2014 in southeast Alaska and the Salish Sea. The ongoing sea star wasting (SSW) epidemic brought...
attention to the condition in sea cucumbers, which has not yet been reported in the peer-reviewed scientific literature. Because this species primarily inhabits subtidal habitats and is generally not as abundant as asteroids, the extent of sea cucumber wasting (SCW) and impacts on population decline is driven by wildlife manager and fisher observations (c.f. extensive citizen science observations of SSW). Wasting was first anecdotally reported in Friday Harbor, WA, in February 2014, subsequently near Admiralty Island (AK) in August 2014, and Santa Catalina Island in September 2014. In April 2015 and again in November 2016, more extensive SCW wasting was observed in the holothurian fishery in southeast Alaska (Tenakee Springs, Chicagof Island) and in the central Salish Sea [44]. In February 2017, SCW was observed at an aquaculture facility in the southern Salish Sea (Manchester, WA) and again in the central Salish Sea and has since been reported more widely in the region (including in the Sechelt and Howe Sound, BC). Anecdotally reported disease signs based on gross observations include non-focal lesions and fissures across the body wall, some sloughing of epidermal tissues, and rapid liquefaction upon collection [44]. The number of individuals recovered at fishing sites anecdotally decreased from 2014 to the present day, suggesting that SCW may have affected overall holothurian population density. Wasting was mostly reported between September and January in 2014 to 2017, which corresponds with seasonal evisceration which is a response to more limited food availability and organ atrophy [45]. During seasonal self-evisceration, viscera are recycled by internal processes [46]. However, self-evisceration does not typically affect body wall tissues, so it is unclear whether SCW is an extension of normal self-evisceration processes. We compared a single wasting A. californicus specimen virome to a single grossly normal A. californicus virome where were collected at the same site (Table 1).

3. Results and Discussion

Twenty-two contiguous sequences met our criteria as matching viral proteins (from here referred to as "viral contigs"), which were primarily within the Picornavirales (Marnaviridae and Dicistroviridae; referred to from here as “picornavirus-like”), Ghabrivirales (Totiviridae; referred to from here as “totivirus-like”), and a single contig matching the Amarillovirales (Flaviviridae; referred to from here as “flavivirus-like”). The Apostichopus californicus wasting library yielded the most viral contigs (Table 1; n=11), followed by Cucumaria miniata and Stichopus horrens (n = 3), and single viral contigs in the Holothuria atra, Holothuria scabra and Holothuria diffcilis libraries. No viral contigs were retrieved from Holothuria pardalis or Synaptula recta libraries. Read recruitment of all libraries against the 22 viral contigs revealed that some were specific to one library, others were more cosmopolitan. Aside from a single picornavirus-like contig (708), all other picornavirus-like contigs were present in only the C. miniata and A. californicus libraries. Conversely, of the 5 totivirus-like contigs, only 2 were specific to a single library, and two (contigs 5835 and 4411) recruited from both Australian and North American holothurians. The flavivirus-like contig (91) recruited from only the North American holothurian libraries, in both C. miniata and A. californicus. Comparing the composition of wasting and grossly normal A. californicus, all contigs except one (picornavirus-like contig 1223) that were assembled in the wasting–affected specimen also recruited reads from the grossly normal specimen library.

These results indicate that most RNA viruses discovered in this survey were restricted to a single, or perhaps sympatric holothurian species, with few that spanned species or genera. Viral host tropism is circumscribed by inherent cell properties, including those affecting entry, replication and release/shedding of virions. Similar species may exhibit greater shared tropism for viruses, which may result in spillover events of pathogens [47]. While variation in factors affecting tropism may result in restricted host range [48], viruses may also infect across highly unrelated hosts, e.g. arboviruses that infect both insects and vertebrates. Amongst the 22 viral contigs retrieved in this survey, genome architecture and features suggest variable hosts, which may include non-holothurian microorganisms and metazoa. Hence, it is not possible to speculate on their host range. However, our data demonstrate that some viral genotypes may have geographically
Table 2. Read recruitment to metagenome-assembled viral genome fragments across all libraries. + indicates that reads recruited from the metavirome library. * indicates wasting-affected individual. Sh = *Stichopus horrens*; Ha = *Holothuria atra*; Hs = *Holothuria scabra*; Hp = *Holothuria pardalis*; Hd = *Holothuria difficilis*; Sr = *Synaptula recta*; Cm = *Cucumaria miniata*; Ac = *Apostichopus californicus*.

<table>
<thead>
<tr>
<th>Species of Metagenome</th>
<th>Contig #</th>
<th>Viral Order</th>
<th>Contig Length (nt)</th>
<th>Holothurian Species Recruited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sh</td>
</tr>
<tr>
<td><em>Stichopus horrens</em></td>
<td>17482</td>
<td>Ghabrivirales</td>
<td>1070</td>
<td>+</td>
</tr>
<tr>
<td><em>Stichopus horrens</em></td>
<td>5835</td>
<td>Ghabrivirales</td>
<td>2224</td>
<td>+</td>
</tr>
<tr>
<td><em>Stichopus horrens</em></td>
<td>10295</td>
<td>Ghabrivirales</td>
<td>796</td>
<td>+</td>
</tr>
<tr>
<td><em>Holothuria atra</em></td>
<td>17799</td>
<td>Picornavirales</td>
<td>793</td>
<td></td>
</tr>
<tr>
<td><em>Holothuria scabra</em></td>
<td>11085</td>
<td>Ghabrivirales</td>
<td>679</td>
<td>+</td>
</tr>
<tr>
<td><em>Holothuria difficilis</em></td>
<td>4411</td>
<td>Ghabrivirales</td>
<td>688</td>
<td>+</td>
</tr>
<tr>
<td><em>Cucumaria miniata</em></td>
<td>4304</td>
<td>Picornavirales</td>
<td>636</td>
<td>+</td>
</tr>
<tr>
<td><em>Cucumaria miniata</em></td>
<td>1913</td>
<td>Picornavirales</td>
<td>1826</td>
<td>+</td>
</tr>
<tr>
<td><em>Cucumaria miniata</em></td>
<td>1781</td>
<td>Picornavirales</td>
<td>782</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>91</td>
<td>Amarillovirales</td>
<td>8883</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>3426</td>
<td>Picornavirales</td>
<td>1158</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>942</td>
<td>Picornavirales</td>
<td>1156</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>792</td>
<td>Picornavirales</td>
<td>767</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>791</td>
<td>Picornavirales</td>
<td>1491</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>708</td>
<td>Picornavirales</td>
<td>1241</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>557</td>
<td>Picornavirales</td>
<td>789</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>45</td>
<td>Picornavirales</td>
<td>1368</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>1580</td>
<td>Picornavirales</td>
<td>854</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>1446</td>
<td>Picornavirales</td>
<td>597</td>
<td>+</td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>1223</td>
<td>Picornavirales</td>
<td>810</td>
<td></td>
</tr>
<tr>
<td><em>Apostichopus californicus</em></td>
<td>1120</td>
<td>Picornavirales</td>
<td>1132</td>
<td>+</td>
</tr>
<tr>
<td>Apostichopus californicus*</td>
<td>1020</td>
<td>Picornavirales</td>
<td>934</td>
<td>+</td>
</tr>
</tbody>
</table>
widespread distribution, since contigs recruited reads from libraries generated from Australian and
North American specimens (Table 2).

**Flavivirus-like genome fragment**: A single flavivirus-like contig (A. californicus contig 91) was
retrieved based on homology with flaviviruses deposited at NCBI. This 8,883 nt contig bore a single
open reading frame (ORF; 8,318 nt; flavivirus polyprotein by CDD search), with 391 nt untranslated
region (UTR) at the 5' end. Within the polyprotein ORF, a region corresponding to the NS5 of
flavivirus RNA-dependent RNA polymerase (RdRp) was observed at the 3' end. Preceding this is a
conserved helicase domain (comprising DEAD N- and C- terminus motifs flanking a Walker motif)
but at the 5' end there was no similarity based on homology with any known flavivirus. However,
alignment of this region by predicted peptide folding revealed significant a significant match to
flavivirus envelope protein (nt positions ~400-700), suggesting it was distant from known
representative flaviviruses. Within this region, there was a slippery sequence frameshift at position
2061 (5' - CUCCUUUUUUUAUC-3') which is also observed in insect-specific flaviviruses [49]. A
start codon (AUG)-flanked hairpin structure, which is characteristic of flaviviruses is at nt positions
232-274 [50, 51] (Fig. 1). There was no similarity to known capsid or membrane proteins at the 3' end
based on CDD search or predicted peptide folding. The first 1000 amino acids contained mostly β
helices, whereas the region closest to the RdRp bore more α-helices, suggesting it may bear the
structural region [52]. The genome arrangement of the A. californicus flavivirus-like genome fragment
is similar to other insect-specific flaviviruses and distinct from flavivirus arboviruses. Reads recruited
from the C. miniata library to this flavivirus-like genome, but not from other holothurian species
(Table 2).

**Figure 1.** Contig map of *Apostichopus californicus* flavivirus-like contig 91. The open reading frame
matched a flavivirus polyprotein by BLASTx [88]. Methyltransferase, NS5 and helicase domains were
identified by comparison against the CDD database at NCBI [41]. The location of the envelope region
was determined by protein folding comparison in Phyre [39]. The hairpin like structure preceding the
Envelope region was determined by folding all sites between start (AUG) codons by mFold [40].

Alignment of the *A. californicus* flavivirus-like genome fragment against closest matches at NCBI
revealed that it was firmly embedded within a clade of insect-specific flaviviruses [53], including
flaviviruses recovered from cephalopods [54] and from invertebrate and vertebrate transcriptomes
[55] (Fig. 2). Enveloped RNA viruses are notable in marine ecosystems because they are not common
constituents of virioplankton and generally experience high decay rates in seawater [56]. Flaviviruses
(+ssRNA) represent important pathogens of mammals, including Dengue Fever, Yellow Fever,
Hepatitis C Virus and Zika Virus. While most flaviviruses recovered to date infect mammals and
represent arboviruses (arthropod-borne viruses) that cause pathology in mammals but do not cause
disease in arthropod vectors, there is an expanding clade of flaviviruses that are found only in
arthropods and never in vertebrates [57]. These viruses have been termed “insect-specific flaviviruses
(iFVs)”. Despite this moniker, iFVs have also recently been observed in teleosts [58, 59], crustaceans
[54] and an elasmobranch [55]. Our discovery of a flavivirus genome fragment associated with a
holothurian host expands the host range attributed to this group, and suggests that these may be
more widespread in invertebrate groups than currently understood. Parry and Asgari [54] inferred
circulation of flaviviruses between vertebrates and invertebrates based on nucleotide frequencies and
active interference RNA response to a shark flavivirus in crabs. The observation of vertebrate-
invertebrate flavivirus associations in marine habitats and in terrestrial habitats may suggest they
have arisen twice. Because flavivirus host range restriction occurs at multiple levels [48], vertebrate-
invertebrate associations have evolved through complex interactions to overcome barriers to
infection and replication. Our observation of a holothurian flavivirus that has greatest similarity to a
shark flavivirus [55] suggests that either these are unique to deuterostomes, or that these may
represent similar invertebrate-vertebrate viral associations.

Figure 2. Phylogenetic representation of Apostichopus californicus flavivirus-like contig 91. The tree was
constructed by performing an alignment of overlapping regions with best BLASTx matches at NCBI using the
CLC Sequence Viewer 8.0 native alignment algorithm. The tree is based on an ~420 amino acid alignment by
neighbor joining and based on Jukes-Cantor distance. Values above nodes indicate bootstrap statistics (>50%)
based on 1000 iterations. The green branches indicate the emerging aquatic and invertebrate-associated
flavivirus clade [54].

Picornavirales-like genome fragments: Sixteen contigs were most similar to Picornavirales (which
comprises picornaviruses, dicistroviruses and iflaviruses) based upon sequence homology to
representatives at NCBI (Fig. 3). Of these, 11 were retrieved from the wasting-affected A. californicus
library, while only one contig was retrieved from the grossly normal A. californicus library. In addition
to A. californicus picornavirus-like genome fragments, 3 fragments were also retrieved from C. miniata.
All contigs bore similarity to RdRp domains of picornavirus polyproteins. Apostichopus californicus
contig 1020 bore 2 overlapping ORFs, which is uncharacteristic of picornaviruses and may indicate
the presence of a frameshift in the overlapping region. Of the 17 picornavirus-like contigs, 11 bore
similarity to rhinovirus capsid (rhv) regions, 6 bore similarity to reverse transcriptase (RT)-like RdRp
domains (one contig, from wasting-affected *A. californicus*, contig 708, contained both rhv and RdRp domains), and one contig bore similarity to a Walker motif (P-loop-NTPase) and to the 3C cysteine protease (picornarin). Of particular note, *Holothuria atra* contig 17799 bore a capsid protein motif most similar to VP4 of Cricket paralysis virus, suggesting that it may belong to the *Dicistroviridae*. Picornaviruses have conserved genome arrangement in the order rhv-helicase-protease-polymerase [60-62]. Interestingly, the one contig bearing both rhv and RdRp motifs was in the arrangement from 5′- to 3′- that was opposite to conserved picornaviruses and iflaviruses and more similar to dicistroviruses [62]. Phylogenetic analyses of detected *Picornavirales* based around the rhv domain demonstrate that they are similar to picorna-like viruses retrieved from transcriptomes of insects, suggesting an invertebrate host (Fig. 4).

![Figure 3](image-url)  
**Figure 3.** Contig maps for *Picornavirales*-like genome fragments recovered from holothurians by viral metagenomics. The color of arrows indicates the expect values of best matches by BLASTx against the non-redundant database at NCBI. Green bars indicate shared homology between contigs based on reciprocal tBLASTx.

![Figure 4](image-url)  
**Figure 4.** Phylogenetic representation of holothurian-associated *Picornavirales*-like genome fragments. The tree was constructed by performing an alignment of an overlapping region (~100 amino acid) of the rhv domain with best BLASTx matches in the non-redundant database at NCBI. The tree was constructed by neighbor joining and based on Jukes-Cantor distance. Values above nodes indicate bootstrap statistics (>50%) based on 1000 iterations.
Picornavirus-like genomes are routinely recovered in invertebrate host-associated and environmental RNA viral metaviromes, [12, 33, 63-70]. There have been no previous reports of picornaviruses associated with any disease in aquatic invertebrates, and their pattern of association between SSW-affected and grossly normal asteroid specimens does not suggest their involvement in disease process [33]. Picornavirales infect a wide range of hosts including metazoa and unicellular eukaryotes [71]. Hence, it is possible that some detected picornavirus-like viruses may infect microscopic eukaryotic constituents of the holothurian microbiome. Our observation of a dicistrovirus-like genome fragment in H. atra may suggests that this infects the species, since the known host range of dicistroviruses includes only arthropods and other invertebrates [62]. Similarly, the phylogenetic similarity of rhv domain-bearing contigs in this survey suggest an invertebrate host rather than unicellular eukaryotic hosts.

**Totivirus-like genome fragments:** Five contiguous sequences matching totivirus genome fragments were recovered from 3 species of holothurian, with most coming from Stichopus horrens (Fig. 5). These represented RdRps (n = 4) and capsid (Cp) proteins (n = 2), with one contig (S. horrens contig 5835) bearing an overlapping Cp-RdRp region, with a possible frameshift at position 882, characteristic of totivirus ORFs [72]. Predicted secondary structure (PSIPRED V4.0 [73]) of the major capsid protein region revealed a 5’ α- and β-strand rich region, followed by a β-helix rich region at the 3’ end, which is most similar to fungal totiviruses and less similar to metazoan totiviruses, which have an α-helix rich 3’ region [74]. The Cp on S. horrens contig 17482, on the other hand, bears an α- and β-rich region at the 5’ end, followed by a β-rich region and terminating in an α-rich region, which suggests it may be more closely related to arthropod totiviruses [74]. Phylogenetically, the Cp regions on both S. horrens 5835 and 17482 match more closely fungal totivirus Cp regions, as do all RdRp fragments recovered in this study (Fig. 6). Hence, these are likely fungal viruses.

**Figure 5.** Contig maps for totivirus-like genome fragments recovered from holothurians. The color of arrows indicates the expect values of best matches by BLASTx [88] to the non-redundant database at NCBI. RdRp = RNA-dependent RNA polymerase, Cp = capsid protein.
The trees were constructed by performing an alignment of an overlapping region of the RdRp (top) and Cp (bottom) domains with best BLASTx matches at NCBI. The trees are based on ~156 amino acid (for RdRp) and 87 amino acid (for Cp) alignments by neighbor joining and based on Jukes-Cantor distance. Values above nodes indicate bootstrap statistics (>50%) based on 1000 iterations.

Fungi do not comprise a large portion of free-living aquatic protistan communities, but they are found in association with invertebrates and are implicated in at least one invertebrate disease [75]. Fungi are easily cultivated from echinoderms, and especially holothurian tissues [76-78], and extracts of holothurian body wall bear antifungal substances [79, 80]. To date there have been no cultivation-independent assessments of holothurian microbiome composition. Because viral metagenomes enrich particles < 0.2 µm, which includes ribosome-sized particles, rRNAs can comprise a large fraction of metaviromic libraries, even after RNAse treatment to remove co-extracted host RNAs. We probed the phylogeny of 18S and 28S rRNAs recovered in RNA viromes, and discovered that most bore signatures of basidiomycete and ascomycete rRNAs (Table 3). Hence, it is not surprising to retrieve fungal viruses in our survey. Nerva et al. [81] observed the diversity of viruses in a fungal collection prepared from Holothuria polii. They discovered that more than half were double-stranded RNA (dsRNA) viruses belonging to the Partitiviridae and Chrysoviridae, the latter of which belongs to the Ghabrivirales. Our observation of putatively fungal totiviruses in echinoderm viromes raises interesting questions about the role of fungi, and interactions between these and their viral predators in host health and ecology.

**Figure 6.** Phylogenetic representations of holothurian-associated totivirus-like genome fragments.

The trees were constructed by performing an alignment of an overlapping region of the RdRp (top) and Cp (bottom) domains with best BLASTx matches at NCBI. The trees are based on ~156 amino acid (for RdRp) and 87 amino acid (for Cp) alignments by neighbor joining and based on Jukes-Cantor distance. Values above nodes indicate bootstrap statistics (>50%) based on 1000 iterations.
Table 3. Taxonomic assignment of contigs matching eukaryotic 18S/28S rRNAs. The number of contigs matching each group is indicated, however should be treated with caution because of un-defined biases in template amplification.

<table>
<thead>
<tr>
<th>Group</th>
<th>Taxonomic Assignment</th>
<th>Holothuria pardalis</th>
<th>Holothuria difficile</th>
<th>Holothuria scabra</th>
<th>Holothuria atra</th>
<th>Synapta recta</th>
<th>Sichopus barnos</th>
<th>Cucumaria miriata</th>
<th>Apostichopus californicus</th>
<th>Apostichopus californicus*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unicellular Eukaryotes</td>
<td>Alveolata</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apusozoa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Amoebozoa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cryptomonada</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heterokonta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metamonada</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Viridiplanta</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>63</td>
<td>18</td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fungi</td>
<td>Ascomycota</td>
<td>17</td>
<td>14</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>17</td>
<td>14</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Basidiomycota</td>
<td>11</td>
<td>8</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>13</td>
<td>14</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Fungi intercae sedis</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metazoa</td>
<td>Annelida</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arthropoda</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Asteroidea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chordata</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>14</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cnidaria</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Echinoidea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Holothuria</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mollusca</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nematoda</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phoronida</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Platyhelminthes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Viral diseases in Holothurians: We expected to see differential representation of wasting-associated viruses in A. californicus which would provide targets for further pathology investigations. However, nearly all viruses discovered in wasting-affected individuals were also present in grossly normal individuals, which is similar to metagenomic analyses performed on sea star wasting-affected individuals [12]. Because of uncertain amplification biases and variable amplification template amounts, it is not possible to quantitatively compare representation between tissue states [16]. Moreover, no challenge experiments with virus-sized material or isolated viruses were performed in
this study. Hence, it would be highly speculative to attribute sea cucumber wasting to any virus with the available information.

Previous work has highlighted several virus-like particles associated with tissues of grossly diseased specimens of the aquacultured Apostichopus japonicus. A skin ulceration and evisceration-associated disease caused mass mortality in the Yellow Sea in 2004-2005. Occluded viral particles were observed by electron microscopy in muscles lining the water vascular system, alimentary canal, connective tissue and respiratory tree, although the virus was not sequenced [82]. In addition, exposure to virus-sized material yielded disease signs in naïve specimens. Separately, [83] reported microscopy-based assessments of acute peristome edema disease in the Yellow Sea in 2007, and described particles with consistent morphology as coronaviruses. Similarly, a spherical virus with a bilayer capsule was described in holothurians experiencing skin ulceration and peristome tumescence disease, and again this appeared to be transmissible [84]. Similar virus particles were also reported in A. japonicus larvae experiencing stomach atrophy, which were identical to virus particles observed in parent gonads [82].

Our observation of a larger number of contigs in wasting A. californicus than a grossly normal specimen is consistent with observations of SSW progression, which showed a general proliferation of viruses in lesion margins [85]. The causes of this viral proliferation, which may exacerbate tissue condition but is likely not the etiological cause of disease [85], may be intracellular hypoxia generated by rapid heterotrophic remineralization in the diffusive boundary layer on respiratory tissues [86]. However, this result could also be a consequence of lower host RNA levels in decaying tissues, which may make more apparent viral genomes that are protected from decay processes [16]. Because SCW occurred simultaneously with seasonal evisceration and in the same season as SSW, it is possible that the mechanism behind SSW (i.e. sequela of boundary layer oxygen diffusion limitation) exacerbated normal apoptotic processes that occur in sea cucumbers as a consequence of food limitation. Determining the etiology of any wildlife disease demands tissue and cell level observations of pathology [87], and rigorous challenge with putative pathogens and environmental conditions. Unfortunately, there have been no reports of SCW histopathologic interrogation to date.

The 8 species of holothurian sampled in this study were broadly similar in ecology. Specimens of A. californicus, S. horrens, H. atra, H. scabra, H. difficilis, H. pardalis, and S. recta are deposit feeders that consume detritus and sediment-bound organisms. A. californicus inhabits subtidal waters in the northeastern Pacific, whereas other species inhabit primarily subtropical and tropical waters of the Indo-Pacific. C. miniata is a suspension feeder that occurs in rocky intertidal zones in the northeastern Pacific. In context of virology, there were no expected differences due to behavior (which may affect contact rates between individuals or with fomites), except for perhaps between C. miniata and other species. However, our observation of read recruits between viral genome fragments detected in C. miniata and A. californicus suggests that geographic location may ultimately influence viral types present in species more strongly than feeding behavior.

Conclusions:

4. Conclusions

To the best of our knowledge, this is the first viral metagenomic survey of holothurians, and our work expands on knowledge of viruses inhabiting echinoderms. Our work revealed the presence of a flavivirus-like genome fragment that falls within an expanding group of aquatic invertebrate flaviviruses, along with novel totivirus and picornavirus-like fragments. While we cannot associate pathology with any of these detected viral genotypes and our comparison is limited to a single individual of sea cucumber wasting, we did not observe a viral genotype that was unique to SCW. Rather, we observed the greatest viral richness in wasting A. californicus, which mirrors viral proliferation during SSW progression. It is recommended that future studies focus on the entry, replication and shedding of novel flaviviruses and the role of fungi and their associated viruses in holothurian ecology.
Author Contributions: Conceptualization, IH and IRT; Methodology, IH, MRJ and IRT; Formal Analysis, IH and MRJ; Resources, IH and IRT; Data Curation, IH and MRJ; Writing – Original Draft Preparation, IH; Writing – Review & Editing, IH, MRJ and IRT; Funding Acquisition, IH

Funding: This work was supported by US National Science Foundation grants OCE-1537111 and OCE-1737127 awarded to IH and an award from the Mario Einaudi Center for International Studies at Cornell University.

Acknowledgments: The authors are grateful to Elliot Jackson and Kalia Bistolas for support of specimen collection and bioinformatic processing. The authors thank Mike Donnellan (Alaska Department of Fish and Game) and Mike Erickson (Alaska Glacier Seafoods) for providing samples of A. californicus from Ketchikan. The Alaska Department of Fish and Game is the permit issuing authority for non-federally managed species in the State waters so the collection A. californicus falls under their existing authority. The authors also thank the Moreton Bay Research Station and Heron Island Research Station for assistance with sample collection. Samples from Moreton Bay were collected under Queensland Parks and Wildlife Service permit it QS2015/MAN316, and from Heron Island under the Great Barrier Reef Marine Park Authority permit G15/37877.1.

Conflicts of Interest: The authors declare no conflict of interest.

References


577 Nyaoke, A. C.; Adkison, M.; Renger, A.; Ng, T. F. F., First isolation of a novel aquatic
578 flavivirus from Chinook Salmon (Oncorhynchus tshawytscha) and its in vivo replication in a
580
581
582
583 60. ICTV, Order - Picornavirales. In Virus Taxonomy, King, A. M. Q.; Adams, M. J.; Carstens, E.
585
586
587
589 C. S. F.; Chan, K.-H.; Yuen, K.-Y., Complete genome analysis of three novel picorna-
591
592
593
595
596
597
599 Dolja, V. V.; Koonin, E. V., Doubling of the known set of RNA viruses by metagenomic
601
602
603
604 64. Sanborn, M. A.; Klein, T. A.; Kim, H.-C.; Fung, C. K.; Figueroa, K. L.; Yang, Y.; Asafo-adjei,
605 E. A.; Jarman, R. G.; Hang, J., Metagenomic analysis reveals three novel and prevalent
606 mosquito viruses from a single pool of Aedes vexans nipponii collected in the Republic of
608
609
610
611 65. Hewson, I.; Bistolas, K. S. I.; Button, J. B.; Jackson, E. W., Occurrence and seasonal dynamics
613
614
615
616 66. Miranda, J. A.; Culley, A. I.; Schvarcz, C. R.; Steward, G. F., RNA viruses as major
617 contributors to Antarctic virioplankton. Environ Microbiol 2016, 18, (11), 3714-3727.
618
619
620
621 67. Culley, A. I.; Lang, A. S.; Suttle, C. A., The complete genomes of three viruses assembled
623
624
625
626 68. Culley, A. I.; Steward, G. F., New Genera of RNA Viruses in subtropical seawater, inferred
628
629
630
631 69. Culley, A. I.; Lang, A. S.; Suttle, C. A., High diversity of unknown picorna-like viruses in
633
634
635
636 70. Culley, A. I.; Lang, A. S.; Suttle, C. A., Metagenomic analysis of coastal RNA virus
638
639
640
641 71. Le Gall, O.; Christian, P.; Fauchet, C. M.; King, A. M. Q.; Knowles, N. J.; Nakashima, N.;
642 Stanway, G.; Gorbalenya, A. E., Picornavirales, a proposed order of positive-sense single-
644
645
646
649
650
651
652 73. Jones, D. T., Protein secondary structure prediction based on position-specific scoring
654
655
656
659 infectious particle of insect-borne totivirus-like Omono River virus has raised ridges and
661
662
663
664 75. Alker, A. P.; Smith, G. W.; Kim, K., Characterization of Aspergillus sydowii (Thom et
666
667
668
669 76. Pivkin, M. V., Filamentous fungi associated with holothurians from the sea of Japan, off the
Zhuravleva, O. I.; Antonov, A. S.; Oleinikova, G. K.; Khudyakova, Y. V.; Popov, R. S.;
Denisenko, V. A.; Pislyagin, E. A.; Chingizova, E. A.; Afiyatullov, S. S.; Virescenosides from
the holothurian-associated fungus Acremonium striatisporum Kmm 4401. Marine Drugs 2019,
17, (11).
Marchese, P.; Garzoli, L.; Gnavi, G.; O’Connell, E.; Bouraoui, A.; Mehiri, M.; Murphy, J. M.;
Varese, G. C., Diversity and bioactivity of fungi associated with the marine sea cucumber
Holothuria poli: disclosing the strains potential for biomedical applications. J Appl Microbiol
Ismail, H.; Lemriss, S.; Ben Aoun, Z.; Mhadhebi, L.; Dellai, A.; Kacem, Y.; Boiron, P.;
Bouraoui, A., Antifungal activity of aqueous and methanolic extracts from the
Adibpour, N.; Nasr, F.; Nematpour, F.; Shakouri, A.; Ameri, A., Antibacterial and
antifungal activity of Holothuria leucospilota isolated from Persian Gulf and Oman Sea.
Nerva, L.; Forgia, M.; Ciuffo, M.; Chitarra, W.; Chiapello, M.; Vallino, M.; Varese, G. C.;
Turina, M., The mycovirome of a fungal collection from the sea cucumber Holothuria polii.
Wang, P.; Chang, Y.; Yu, J.; Li, C.; Xu, G., Acute peristome edema disease in juvenile and
adult sea cucumbers Apostichopus japonicus (Selenka) reared in North China. J Invert Pathol
of the pathogens associated with skin ulceration and peristome tumeescence in cultured sea
Aquino, C. A.; Besemer, R. M.; DeRito, C. M.; Kocian, J.; Porter, I. R.; Raimondi, P. T.; Rede,
J. E.; Schiebelhut, L. M.; Sparks, J. P.; Wares, J. P.; Hewson, I., Evidence for boundary layer
oxygen diffusion limitation as a key driver of asteroid wasting. BioRxiv 2020, doi:
https://doi.org/10.1101/2020.07.31.231365
Frakolaki, E.; Kaimou, P.; Moraiti, M.; Kalliampakou, K. I.; Karampetsou, K.; Dotsika, E.;
Liakos, P.; Vasilacopoulou, D.; Mavromara, P.; Bartenschlager, R.; Vasilaki, N., The role of
tissue oxygen tension in Dengue virus replication. Cells 2018, 7, (12).
Work, T. M.; Aeby, G. S., Systematically describing gross lesions in corals. Dis Aquat Organ
2006, 70, (1-2), 155-60.
Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J.,
Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.