Diversity of sea star-associated densoviruses and transcribed endogenized viral elements of densovirus origin

Elliot W. Jackson[1*], Roland C. Wilhelm[2], Mitchell R. Johnson[1], Holly L. Lutz[3,4], Isabelle Danforth[4], Joseph K. Gaydos[5], Michael W. Hart[6], Ian Hewson[1]

[1] Department of Microbiology, Cornell University, Ithaca NY USA
[2] School of Integrative Plant Sciences, Bradfield Hall, Cornell University, Ithaca, NY USA
[3] Department of Pediatrics, School of Medicine, University of California San Diego, La Jolla, California, USA
[4] Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA
[5] SeaDoc Society, UC Davis Karen C. Drayer Wildlife Health Center – Orcas Island Office, Eastsound, WA 98245 USA
[6] Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

* Correspondence: ewj34@cornell.edu; Tel.: +1-2 31-838-6042

Journal of Virology

Abstract:

A viral etiology of Sea Star Wasting Syndrome (SSWS) was originally explored with virus-sized material challenge experiments, field surveys, and metagenomics leading to the conclusion that a densovirus is the predominant DNA virus associated with this syndrome, and, thus, the most promising viral candidate pathogen. Single-stranded DNA viruses are however highly diverse and pervasive among eukaryotic organisms which we hypothesize may confound the association between densoviruses and SSWS in sea stars. To test this hypothesis and assess the association of densoviruses to SSWS, we compiled past metagenomic data with new metagenomic-derived viral genomes from sea stars collected from Antarctica, California, Washington, and Alaska. We used 179 publicly available sea star transcriptomes to complement our approaches for densovirus discovery. Lastly, we focus the study to SSaDV, the first sea star densovirus discovered, by documenting its biogeography and putative tissue tropism. Transcriptomes contained mostly endogenized densovirus

elements similar to the NS1 gene, while numerous complete and near-complete densoviral genomes were recovered from viral metagenomes. SSaDV was associated with nearly all tested species from southern California to Alaska, and in contrast to previous work, we show SSaDV is one genotype among a high diversity of densoviruses present in sea stars across the west coast of the United States and globally that are commonly associated with grossly normal (i.e. healthy or asymptomatic) animals. The diversity and ubiquity of these viruses in wild sea stars confounds the original hypothesis that one densovirus was the etiologic agent of SSWD.

Importance:

The primary interest in sea star densoviruses, specifically SSaDV, has been their association with Sea Star Wasting Syndrome (SSWS), a disease that has decimated sea star populations across the west coast of the United States since 2013. The association of SSaDV to SSWS was originally drawn from metagenomic analysis, which was further studied through field surveys using qPCR, concluding that that it was the most likely viral candidate in the metagenomic data based on its representation in symptomatic sea stars compared to asymptomatic specimens. We reexamined the original metagenomic data with additional genomic datasets and found that SSaDV was one of ten densoviruses present in the original dataset and was no more represented in symptomatic sea stars than in asymptomatic sea stars. Instead, SSaDV appears to be a widespread, generalist virus that exists among a large diversity of densoviruses present in sea star populations.

Introduction:

Single-stranded (ss) DNA viruses are among the most diverse and prevalent group of viruses infecting eukaryotes, bacteria, and archaea (1–4). Recognition of their ubiquity has been made possible through the use of rolling circle amplification that preferentially amplifies circular nucleic acid templates prior to high-throughput sequencing (5, 6). As a

result, ssDNA viruses that possess circular genomes are significantly overrepresented compared to those with linear genomes. There are currently nine established families of ssDNA viruses that infect eukaryotes, only two of which possess linear genomes - *Bidnaviridae* and *Parvoviridae* (7). While our knowledge of the circular ssDNA viruses has expanded tremendously, the discovery of linear ssDNA viruses has lagged. Collectively, the known viral diversity of these two families likely represents only a small proportion of actual extant diversity, particularly within the subfamily *Densovirinae* (family *Parvoviridae*).

Currently 17 viral species are recognized by the International Committee on the Taxonomy of Viruses that belong to the subfamily *Densovirinae* (commonly referred to as densoviruses) (8). Densoviruses infect invertebrates and until recently, were only known to infect insects (orders Blattodea, Diptera, Hemiptera, Hymenoptera, Lepidotera, and Orthoptera) and decapod crustaceans (shrimp and crayfish) (9). Unlike circular ssDNA virus groups, the discovery of novel linear densoviruses genomes has occurred primarily via the use of classical methods, such as viral purification through cell culture or viral enrichment via an animal model followed by nucleic acid sequencing, but these classical methods are more recently supplemented by the use of high-throughput sequencing to explore viral diversity. Transcriptomic and metagenomic analyses have generated a growing body of evidence to suggest that densoviruses infect a more phylogenetically diverse array of invertebrate hosts outside of the phylum Arthropoda (10–13). Expanding the known potential host range of these viruses will lead to insights into biology and evolution of densoviruses and the hosts they infect.

The 2014 discovery of densoviruses associated with sea stars and sea urchins (phylum: Echinodermata) was a significant step towards expanding the range of densoviruses beyond arthropods (11, 13, 14). The primary interest in echinoderm densoviruses has been their association with Sea Star Wasting Syndrome (SSWS; also referred to as Sea Star Wasting Disease and Asteroid Idiopathic Wasting Syndrome) which is an epidemic affecting sea stars on the east and west coast of North America (11, 15). The gross signs of SSWS range from lethargy and limb curling, loss in turgor, epidermal ulceration and tissue loss, ray autotomy, and eversion of viscera through the body wall all of which generally leads to the

74 animal's death (11, 16). There is currently no clinical pathological case definition for SSWS so a symptomatic individual is
75 defined by the presentation of these gross signs (16). The etiology of SSWS is also currently unknown. A densovirus,
76 commonly known as sea star associated densovirus or SSaDV, was hypothesized to cause SSWS but it has yet to be
77 determined if SSaDV is a pathogen (11, 17). Originally, SSaDV was thought to be the etiological agent of SSWS observed in
78 both Pacific and Atlantic sea star populations (11, 15). However, the recent discovery of a genetically similar second
79 densovirus associated with sea stars on the Atlantic coast, *Asterias forbesi* associated densovirus or AfaDV, raises new
80 questions about densovirus diversity and the putative role of densoviruses in SSWS (13). The discovery of AfaDV
81 prompted us to explore the diversity of densoviruses at a greater geographic scale and to define the biogeography of
82 SSaDV and expand our understanding of the ecology, evolution, and diversity of these viruses.
83 Here, we employ a multi-omic approach to document the biodiversity of densovirus populations in sea stars by first
84 reassessing the original metagenomic dataset leading to the association of densoviruses and SSWS. We used publicly
85 available sea star transcriptomes/genomes and sea star viral metagenomes for viral discovery, and PCR to document the
86 prevalence, putative tissue tropism, and biogeography of SSaDV. We report the discovery of >30 novel sea star
87 densoviruses associated with sea stars from the Southern Ocean around Antarctica (11 genomes) and from the temperate
88 eastern Pacific (24 genomes), and the observation of numerous endogenized viral elements (EVEs) from sea star
89 transcriptomes and genomes. We found that SSaDV putatively has a wide tissue tropism and is associated with sea stars in
90 the eastern Pacific across a broad latitudinal range from southern California to Alaska, corroborating previous findings by
91 Hewson *et al.*, 2014 (11). The identification of SSaDV as one among many densoviruses infecting sea stars on the west
92 coast of the United States suggests that densoviruses may comprise a normal component of the sea star microbiome,
93 bringing into question the association of densoviruses to SSWS.
94
95 **Materials and Methods:**
96

Tissue Collection & DNA Extractions:

We collected 887 tissue samples from 660 individual sea stars spanning 12 species from 42 locations from the temperate eastern Pacific coast of the United States from 2005, 2014-2019 (**Supplemental Table 1, Supplemental Table 2, Supplemental Table 3**). Thirty samples were collected during the peak of SSWS epidemic observed from mid 2013-2015 in the Northeast Pacific (18). The majority of samples collected were from asymptomatic individuals (85% of sea stars sampled). Tissues were collected from sea star specimens either by vivisection immediately upon collection followed by flash freezing in liquid nitrogen and storage at -80°C or -20°C until dissection, or sampled from individuals in the field non-lethally and preserved in RNAlater (SigmaAldrich) or EtOH (**Supplemental Table 1**). Coelomic fluid samples were collected only from vivisected specimens using a 25G x 1. (0.5mm x 25mm) needle attached to a 3 mL syringe inserted through the body wall into the coelomic cavity. DNA was extracted from tissues and coelomic fluid using the Zymo Research Quick-DNA Miniprep Plus kit or the Zymo Research Duet DNA/RNA Miniprep kit following the manufacture's protocol. DNA was quantified using a Nanodrop or Quant-iT PicoGreen dsDNA Assay kit (Invitrogen) (**Supplemental Table 1, Supplemental Table 3**).

PCR and Sanger Sequencing of SSaDV:

Primers were designed targeting the structural gene of SSaDV using Primer3web (version 4.1.0). The primers were as follows: VP1 forward primer (5'- TGGCCACTCATCATGTCTCT -3') and VP1 reverse primer (5' – CTTGGGGTCCTTCATGAGC – 3'). NEB Q5 High-Fidelity DNA Polymerase was used following the manufacturer's protocol for a 50µl reaction volume. Thermal cycling was performed in a Bio-Rad C1000™ Thermal Cycler using the following conditions: initial denaturing at 98°C for 30 seconds followed by 35 cycles of denaturing (98°C for 10 seconds), annealing (67°C for 20 seconds), and extension (72°C for 20 seconds) followed by a final extension of 72°C for 2 minutes. Annealing temperature was based on NEB Tm Calculator recommendation using default primer concentration of 500nM.

The resulting amplicon of the PCR reaction was 534 nucleotides (nt). All PCR reactions included a positive control, a kit negative control, and PCR reagent negative control to account for false positives and false negatives. 10-15µl of a PCR reaction was used for gel visualization. The remaining PCR product was processed using a ZR-96 DNA Clean & Concentrator-5 kit (Zymo Research) and submitted to Cornell Core Biotechnology Resource Center Genomics Facility for DNA sequencing. DNA sequencing was performed on Applied Biosystems Automated 3730xl DNA Analyzers using Big Dye Terminator chemistry and AmpliTaq-FS DNA Polymerase.

Viral Metagenome Preparation:

Five RNA viral metagenomes were prepared from five different species of sea stars - *Pisaster ochraceus* (n=1), *Labidiaster annulatus* (n=1), *Leptasterias* spp. (n=1), *Mediaster aequalis* (n=1), and *Neosmilaster georgiansus* (n=1) (**Supplemental Table 4**). The preparation of viral metagenomes followed Hewson *et al.,* 2018 modified from Thurber *et al.*, 2009 (17, 19). Pyloric caeca from sea stars were homogenized in a 10% bleach-cleaned NutriBullet with 0.02µm-filtered 1X PBS. Tissue homogenates were pelleted by centrifugation at 3,000 x g for 5 minutes, and the supernatant was syringe filtered through Millipore Sterivex-GP 0.22µm polyethersulfone filters into 10% bleach-treated and autoclaved Nalgene Oak Ridge High-Speed Centrifugation Tubes. Filtered homogenates were added to a 10% (wt/vol) PEG-8000 in 0.02µm-filtered 1X PBS with a final volume of 35mL and precipitated for 20 hours at 4°C. Precipitated nucleic acids, and other cell material, were pelleted by centrifugation at 15,000 x g for 30 minutes. The supernatant was decanted and pellets were resuspended in 2mL of 0.02µm-filtered 1X PBS. Half of the sample (1mL) was treated with 0.2 volumes (200µl) of $CHCl_3$, inverted three times and incubated at room temperature for 10 minutes. After a brief centrifugation, 800µl of supernatant was transferred into 1.5mL microcentrifuge tube. Samples were treated with 1.5µl of TURBO DNase (2U/µl) (Invitrogen), 1µl of RNase One (10U/µl) (Thermo Scientific), and 1µl of Benzonase Nuclease (250U/µl) (MilliporeSigma) and incubated at 37 °C for 3 hours. 0.2 volumes (160µl) of 100mM EDTA was added to the sample after incubation. Viral

RNA was extracted using the ZR Viral RNA kit (Zymo Research). RNA was converted into cDNA and amplified using the WTA2 kit (Sigma-Aldrich). Prior to sequencing, samples were processed using a ZR-DNA Clean & Concentrator-5 kit, and DNA was quantified by Quant-iT PicoGreen dsDNA Assay Kit. Samples were prepared for Illumina sequencing using the Nextera XT DNA library preparation kit prior to 2x250bp paired-end Illumina MiSeq sequencing at the Cornell Core Biotechnology Resource Center Genomics Facility.

Metagenome-derived Viral Genome Discovery

Raw paired-end reads were quality trimmed to remove Illumina adapters and phiX contamination. Reads were merged and normalized to a target depth of 100 and a minimum depth of 1 with an error correction parameter. Read quality filtering, trimming, contamination removal, merging, normalization, and read mapping were performed using the BBtools suite (20). Both merged and unmerged reads were used for *de novo* assembly using the default parameters excluding the read error correction option in SPAdes v 3.11.1 (21). Contigs shorter than 3000 nt were discarded after assembly and the remaining contigs were subjected to tBLASTx against a curated in-house database containing 453 genomes from all nine families of eukaryotic ssDNA viruses. Contigs with significant sequence similarity at *e*-value < $1x10^{-8}$ to a densovirus genome were reviewed in Geneious version 9.1.5 (22). Open reading frames (ORFs) were called in Geneious using a minimum size of 550 nt with a standard genetic code and a start codon of ATG. Hairpin structures were identified using Mfold (23). After verifying the contigs as densovirus sequences, reads were mapped back to contigs with a minimum identity of 0.95 to obtain average read coverage and total reads mapped to contigs (**Supplemental Table 5**). All densovirus sequences have been deposited in GenBank (**Table 1**). In addition to the 5 viral metagenome libraries sequenced in this study, we reanalyzed 30 DNA and 21 RNA viral metagenomes published elsewhere (NCBI BioProjects: PRJNA253121, PRJNA417963, PRJNA637333) using the assembly approach described above (**Supplemental Table 4**) (11, 17).

Sea Star Transcriptome Analysis:

A total of 179 sea star RNA-seq paired-end libraries were downloaded from NCBI (**Supplemental Table 6**). FastX was used to remove reads with lengths <50 nt and a quality score <30 (24). Trimmomatic was used to trim adapters (25). Libraries were assembled using default parameters in Trinity v2.1.1 (26). Assembled contigs were annotated against a protein database of sea star densoviruses using DIAMOND with an e-value cutoff of $<1 \times 10^{-5}$ (27). Contigs containing sequences with significant similarity to the sea star associated densovirus protein database were isolated, ORFs were called, and amino acid sequences from ORFs were further checked by BLASTp against the NCBI non-redundant database. Top BLAST results for each densovirus-like sequence against the NCBI non-redundant database were downloaded for amino acid MUSCLE alignment and visualized using Geneious (28, 22). To verify if the sequences were from the host, BLASTn was performed querying isolated densovirus-like sequences against available sea star genomes (*Asterias rubens*, NCBI taxid: 7604, and *Acanthaster planci*, 133434).

Densovirus Phylogenetics:

Phylogenetic analysis was performed on 89 densoviruses sequences that included 39 sea star-associated densoviruses and 50 densovirus sequences from complete or near-complete genomes available on NCBI. All densovirus sequences included in the phylogeny were from extant viruses (i.e., no endogenized densovirus elements). Amino acid sequences from the NS1 gene were aligned with MUSCLE using default parameters (28). The region of NS1 used for alignment (sequence length of 434.9 ± 40.2 (mean $\pm$ standard deviation)) spanned motif I of the replication initiation motifs past Walker C of the Walker box ATPase motifs. Phylogenetic relationships between densovirus genomes were inferred by a LG + G + I + F substitution model selected by smart model selection (SMS) in PhyML 3.0 (29). Branch support was determined by bootstrapping for 100 iterations. The resulting maximum likelihood phylogenetic tree was

visualized and annotated using iTOL (30). CD-HIT was used to identify viral species using a 85% amino acid sequence identity of NS1 (31, 32).

Data availability:

Densovirus genome sequences have been deposited in GenBank under accession numbers MT733013 to MT733051 (**Table 1**). Raw viral metagenomic sequence data have been deposited under BioProject numbers PRJNA253121, PRJNA417963, and PRJNA637333 (**Supplemental Table 4**). Assembled sea star transcriptome data have been deposited in OSF (https://osf.io/bh8cr/?view_only=928795303627432 9a89270b8c8a51151) (**Supplemental Table 6**).

Results:

Reanalysis of metagenomes published in Hewson *et al.,* 2014:

The reanalysis of the viral metagenomic data presented in Hewson *et al*., 2014 (11), led to the discovery of 9 additional densovirus genomes in addition to SSaDV (**Figure 1**, **Table 1**). The densovirus contigs ranged in nucleotide length from 3391 to 6053 nt (5002 ± 921 (mean ± standard deviation)). SSaDV was the only densovirus assembled into a complete or near complete genome across multiple metagenomes and had the highest read recruitment among all libraries (**Table 1**, **Supplemental Table 5**). The previously published partial SSaDV genome (5,050 nt) lacked the NS3 ORF and inverted terminal repeats (ITRs) (11). In this study, we recovered three SSaDV genomes of varying sizes from three of the 32 metagenomes (**Table 1**). The largest of these genomes (6,053 nt) contained the expected ORFs (NS1, NS2, NS3, and VP), ITRs, and hairpins within the ITRs, and therefore likely represents a complete genome. It is possible that the ITR region of the genome is not complete, due to challenges posed by assembling regions with high frequency of repeats using short read technology. Members of the genus *Ambidensovirus* typically have ITRs >500 nt, which is considerably longer

than the ITR regions we observed (9). The length of the ITRs in SSaDV were 260 nt on both sides of the genome and contained canonical hairpin structures that were 223 nt and are thermodynamically favorable (ΔG = -106.40).

SSaDV biogeography and tissue tropism:

A total of 148 of 660 animals were virus positive for SSaDV based on the PCR assay, equating to a global prevalence of 22.4% (**Figure 2, Supplemental Table 1**). No samples were PCR positive from tissues collected in 2005. A total of 126 of 148 PCR amplicons were successfully Sanger sequenced, all confirming the specificity of the PCR assay (**Supplemental Table 1)**. Only 22 of 102 (21.6%) symptomatic (i.e. SSWS affected) sea stars were PCR positive, and 126 of 558 (22.6%) asymptomatic sea stars tested PCR positive. SSaDV was detected in 25 of the 42 locations that spanned a broad latitudinal range in the eastern Pacific from southern California to southeastern Alaska (**Figure 2, Supplemental Table 2)**. Nine of 12 sea star species tested positive, which include the following species (virus positive / sample total) : *Pisaster ochraceus* (87/287), *Pisaster brevispinus* (10/10), *Pisaster giganteus* (3/9), *Pycnopodia helianthoides* (4/72), *Evasterias troschelii* (26/100), *Dermasterias imbricata* (8/34), *Henricia spp.* (3/18), *Leptasterias spp.* (6/42), and *Patiria miniata* (1/85). Only one individual was tested for each of the three species in which SSaDV was not detected. These included *Orthasterias koehleri*, *Pteraster tesselatus*, and *Solaster stimpsoni*.

Fine dissections of *Pisaster ochraceus* (n= 26 individuals), *Evasterias troschelii* (n = 11), and *Pisaster brevispinus* (n = 10) collected from Langley Harbor, Washington were used assess putative tissue tropism. The viral prevalence among tissues was calculated by the number of tissues positive divided by the total number of tissues collected between these three species. SSaDV was detected most frequently in the pyloric caeca (89%, 40/45) followed by tube feet (36%, 17/47), stomach (11%, 5/46), body wall (11%, 5/47), and gonads (10%, 4/42) (**Figure 2, Supplemental Table 2**). Similar to AfaDV, SSaDV was not detected in the coelomic fluid (0%, 0/47) (13).

Genome discovery, genome comparison, motif annotation, and phylogeny:

An additional 29 densovirus genomes were recovered from newly prepared and reanalyzed sea star metagenomes (**Table 1).** The densovirus contigs ranged in size from 3061 to 5963 nt (5179.0 ± 684.4 (mean ± standard deviation Most densovirus-containing contigs (n = 28) corresponded to near complete genomes containing all the expected ORFs but lacking either ITRs or hairpins with the ITRs. The average size of the ORFs found in sea star densovirus were as follows (± standard deviation): NS1 1694.8 nt (± 33.3), NS2 883.1 nt (± 28.5), NS3 807.8 (± 109.2), and VP 2723.0 nt (± 98.2). The pairwise nucleotide identity was greater than amino acid (aa) identity among sea star densovirus genomes for NS1, NS3, and VP ORFs (**Figure 3, Supplemental Figure 1**). The NS1 ORF had the highest sequence conservation (55.7% average nt and 43.2% average aa pairwise identity) compared to NS3 (32.6% average nt and 18.6% aa acid pairwise identity), and VP (43.8% average nt and 34.2% average aa pairwise identity). The current delineation for a new parvovirus species is based on the pairwise amino acid sequence identity of NS1. Parvoviruses encoding for NS1 proteins with a >85% pairwise amino acid sequence identity are considered the same viral species (32). Using this species definition, 29 new sea star densovirus species were defined from the 39 genotypes discovered. There were 8 viral species that contained 2 or 3 genotypes, and 21 species contained a single genotype. (**Table 1**).

All sea star densoviruses discovered thus far have ambisense genomes that fall into subgroups A and B, which differ only by the VP ORF organization (33) (**Figure 4**). The NS1 and VP ORFs identified in this study contain all the expected motifs that are characteristic of densoviruses (34). These motifs include: RCR I and RCR II of the replication initiation motifs, Walker A, B, and C of the NTP-binding and helicase motifs, and the viral phospholipase A_2 motif.

Identification of EVEs of densovirus origin

A total of 8 of the 179 transcriptomic libraries contained contigs with densovirus-like sequences. Ten densovirus-like sequences were found among the 8 libraries based on homology searches against the sea star-associated densovirus database. These densovirus-like sequences only encoded part of NS1, lacked RCR motifs, and were not the typical coding

length found in sea star-associated densovirus genomes. These sequences are likely endogenized viral elements of densovirus origin due to the fragmentation of the viral genome and the missing enzymatic motifs. The endogenized densovirus elements primarily contained Walker box ATPase motifs with homology to parvoviruses and densoviruses from a broad diversity of hosts (**Figure 5, Supplemental Table 7**). The transcriptomes containing densovirus-like sequences came from the following species: *Acanthaster planci* (SRA run ID: DRR072325), *Patiria pectinifera* (SRR5229427), *Echinaster spinulosus* (SRR1139455 and SRR2844624), *Acanthaster brevispinus* (SRR276461), *Linckia laevigata* (SRR5438553), and *Asterias rubens* (SRR1139190 and SRR3087891) (**Supplemental Table 6**).

Discussion:

The initial investigation for a viral agent associated with SSWS performed viral metagenomic surveys (DNA and RNA) to compare the viral consortia between species and between asymptomatic and symptomatic individuals to determine the most likely viral candidate for further investigation (11). The conclusions from these molecular surveys was that SSaDV was more represented in metagenomic libraries of symptomatic individuals and was present in symptomatic metagenomes prepared from multiple sea star species (11). However, these data show that SSaDV was one of ten densoviruses present, and was neither more abundant by number of reads per library comparing asymptomatic to symptomatic individuals (based on read mapping analysis) nor more prevalent between libraries comparing symptomatic to asymptomatic individuals (**Figure 1, Supplemental Table 5**). This result contradicts the original conclusion from the metagenomic data that SSaDV was associated with SSWS. The difference in results is directly attributable to the difference in the bioinformatic assembly approach. The original analysis took an overlap-layout-consensus global assembly approach using the 28 DNA viral metagenomes not including the 4 RNA viral metagenomes (11). In this study, we choose SPAdes, a more sensitive de Bruijn graph assembler, and included the RNA viral metagenomes in the analysis which contained six of the nine novel densoviruses found in this dataset (**Table 1**). We did find SSaDV to be, on average, the most abundant densovirus by read mapping analysis and, thus, the most consistently assembled, likely biasing its

assembly and discovery in the original analysis. The higher representation of SSaDV among metagenomic libraries may be the result of higher viral enrichment in those samples prior to sequencing rather than a result of greater viral loads prior to metagenomic preparation. While a higher abundance of SSaDV could reflect important host-virus biology, it remains to be determined whether greater viral loads, measured by qPCR, of SSaDV has any biological significance. A pitfall of viral metagenomics from animal tissue using a viral enrichment method is the significant variability in non-viral genetic material between viromes within a study making quantitative comparisons difficult (35). According to our metagenomic survey, SSaDV is one species within a diverse extant population of densoviruses present in sea star populations on the west coast of the United States.

The discovery of densoviruses associated with sea stars collected from China, Antarctica, and the Pacific and Atlantic coasts of the United States indicates their ubiquitous distribution and substantial extant diversity (**Table 1, Figure 3, Supplemental Figure 1**). The diversity observed in this study is likely a small fraction of the total diversity among echinoderms, considering these viruses have also been found in sea urchins (14). Sea star associated densoviruses also seem to be pervasive in wild populations. The two densovirus genotypes with the best-described ecological characteristics, SSaDV and AfaDV, share striking similarities. Both viral genotypes are not species-specific, found across a large geographic range, are commonly found in asymptomatic individuals, and have a wide tissue tropism with pyloric caeca being the primary tissue of detection (**Figure 2**) (13). This set of characteristics suggests that both viruses form persistent infections in sea stars.

The genus *Ambidensovirus*, to which both previously described sea star associated densoviruses belonged to based on genome organization, was recently divided into seven newly proposed genera to resolve paraphyly within the genus (8). In this new arrangement, SSaDV and CqDV (*Cherax quadricarinatus* (shrimp) densovirus, the most genetically similar densovirus to SSaDV prior to the discovery of AfaDV) were assigned to the genus *Aquambidensovirus,* putatively uniting all aquatic densoviruses (36, 13, 8). Our phylogenetic analysis did not support the monophyly of sea star associated densoviruses within the newly proposed *Aquambidensovirus* genus, nor did all aquatic densoviruses cluster into a single

well-supported clade (**Figure 3, Supplemental Figure 1**). Newly proposed classification schemes within the *Ambidensovirus* genus would greatly benefit from the inclusion of broader taxonomic sampling before proposing new systematic arrangements of this highly diverse genus.

Given the lack of immortal cell-cultures, the discovery of echinoderm densoviruses has been primarily through metagenomics, and that constraint was the motivation for our analysis of transcriptomes as an additional and alternative option for densovirus discovery. Host transcriptomes have been a rich source of viral discovery from eukaryotes and have expanded our knowledge of host associations for many viral groups including parvoviruses (12, 37, 38). However, we did not find transcriptomes to be an effective method for the purpose of densovirus discovery compared to viral metagenomes particularly RNA viral metagenomes. This could be due to various methodological reasons. First, the viral metagenomes prepared for this study were enriched for encapsulated nucleic acids with a cDNA enrichment step; by contrast, transcriptomes target mRNA through rRNA depletion and/or through selection for poly-A tails. Second, to detect DNA viruses from a host transcriptome requires tissue containing an active infection, which may not be detectable without very high sequence depth. The transcribed EVEs found in this study were only detected in transcriptomes larger than 2.4 Gbases. Third, the genomes discovered in the RNA viral metagenomes are likely ssDNA that was carried through the RNA extraction process. ssDNA is an uncommon nucleic acid template for non-viral material and a difficult template to remove during RNA extraction. Most commercial kits use DNases that preferentially target dsDNA and inefficiently cleave ssDNA. Without preferentially targeting mRNA prior to cDNA synthesis in addition to enriching for encapsulated nucleic acid, the chances of picking up ssDNA in a pool of RNA is much higher.

None of the transcriptome-derived densovirus-like sequences appeared to be extant densoviruses based on ORF architecture and motif repertoire. We conclude that these densovirus-like sequences are likely transcribed EVEs present in host cells. EVEs from *Asterias rubens* and *Acanthaster planci* could be traced to their genomes, while our inability to trace others reflects the lack of publicly available host genomes. The putative EVEs present in *Asterias rubens* were nearly identical to those previously reported in the same host (39) though most we observed had low sequence identity to

previously identified EVEs from other invertebrates. It is likely that these EVEs have been established in the germline of *Asterias rubens*, and our findings corroborate previous work proposing sea star densoviruses can infect germ line cells (13). The expression of these EVEs in *Asterias rubens* was found to trigger the RNA interference (RNAi) response, specifically the Piwi-dependent pathway, signifying these EVEs are still recognized as foreign and are regulated through the immune system (39). This RNAi response has been widely observed in terrestrial invertebrate genomes containing EVEs descending from densovirus (40). We expect that the expansion of echinoderm genomes, and corresponding small RNA libraries, will further support this conclusion.

Essentially all observed EVEs retained the Walker Box ATPase motifs which collectively function as a helicase (**Figure 5**) (41). This helicase domain belongs to the superfamily III helicases (SF3) which are more broadly grouped as AAA + ATPases (42). SF3 helicases are only encoded by DNA and RNA viruses so their presence in cellular genomes must be the result of endogenization (43, 44). The retention of the Walker box ATPase motifs among endogenized densovirus elements has been observed across a diverse range of invertebrate hosts, suggesting a beneficial function for coopting and possibly maintaining the function of the SF3 helicase (10, 12, 40, 45, 46). The adaptive benefit of a Walker Box ATPase-containing EVE has been demonstrated in the pea aphid (*Acyrthosiphon pisum*), where wing development was regulated by two modified densovirus NS1 EVEs, which only retained the Walker Box ATPase motifs (46). The expression of these two EVEs in crowded conditions initiated wing development, which could be suppressed by knocking down their expression. These results demonstrate that this viral gene can be co-opted by the host to modulate the response of a phenotypically plastic trait to environmental cues. Another plausible hypothesis for EVE function is the ability to enhance or prime the immune system against new infections (40). However, we observed little sequence identity between extant sea star-associated densoviruses and the EVEs in their transcriptomes. These sequence differences suggest that these EVES are unlikely to have a role in priming the piRNA response against new infections.

We employed metagenomic and transcriptomic approaches to explore the diversity of sea stars associated-densoviruses, while advancing understanding of the biogeography of SSaDV, the first densovirus found in sea stars.

Empirically, we found that viral metagenomes provided a more effective resource for densovirus discovery compared to host transcriptomes. We discovered 37 new densovirus genomes from sea stars and identified EVEs expressed in host transcriptomes that are of densovirus origin based on detection of the tripartite SF3 helicase domain in these EVEs. Using PCR, we found SSaDV to have a putatively wide tissue tropism, with the pyloric caeca being the most consistent tissue for viral detection. SSaDV was detected across a broad latitudinal range in the northeastern Pacific from southern California to Alaska and found in tissues in nearly all sea star species tested. These results corroborate the hypothesis that these viruses are common among populations and suggest they form persistent infections in sea stars. Given the diversity of densoviruses and their broad distribution among tissues, populations, and species of both asymptomatic and symptomatic sea stars, we propose that the association of SSaDV with Sea Star Wasting Syndrome should be critically reassessed relative to the mounting evidence that this virus may not be pathogen that causes this disease and instead a common constituent of these animals' microbiomes.

Acknowledgments

The authors thank Dr. Sarah Gravem (Oregon State University), Dr. Bruce Menge (Oregon State University), Dr. Lauren Schiebelhut (UC Davis), Dr. Michael Dawson (UC Merced), Elizabeth Ashley (SeaDoc Society), Erika Nilson (SeaDoc Society) , Dr. Melissa Pespeni (University of Vermont), Dr. John Ware (University of Georgia), Christoph Pierre (UC Santa Barbara), Carter Urnes (National Park Service), Dr. Steven Fradkin (National Park Service), Dr. Nathalie Oulhen (Brown University), Dr. Gary Wessel (Brown University), Dr. Sarah Cohen (San Francisco State University), Dr. Margaret Amsler (University of Alabama), Dr. James McClintock (University of Alabama), Dr. Chuck Amsler University of Alabama), and Sean Williams (Hoonah Indian Association) for assistance in sample collection.

This work was supported by NSF grants OCE-1537111 and OCE-1737127 awarded to IH, USGS contract G19AC00434 awarded to IH and T. Work, and NSF grant PLR-1341333 awarded to CD Amsler and JB McClintock. This work was also supported by the Cornell Atkinson Center's Sustainable Biodiversity Fund and Andrew W. Mellon Student Research Grant awarded to EWJ.

References

1. Labonté JM, Suttle CA. 2013. Previously unknown and highly divergent ssDNA viruses populate the oceans. ISME J 7:2169–2177.

2. Zhao L, Rosario K, Breitbart M, Duffy S. 2019. Eukaryotic circular rep-encoding single-stranded DNA (CRESS DNA) viruses: ubiquitous viruses with small genomes and a diverse host range, p. 71–133. *In* Advances in virus research. Elsevier.

3. Roux S, Krupovic M, Daly RA, Borges AL, Nayfach S, Schulz F, Sharrar A, Carnevali PBM, Cheng J-F, Ivanova NN. 2019. Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. Nat Microbiol 4:1895–1906.

4. Tisza MJ, Pastrana DV, Welch NL, Stewart B, Peretti A, Starrett GJ, Pang Y-YS, Krishnamurthy SR, Pesavento PA, McDermott DH. 2020. Discovery of several thousand highly diverse circular DNA viruses. Elife 9:e51971.

5. Dean FB, Nelson JR, Giesler TL, Lasken RS. 2001. Rapid amplification of plasmid and phage DNA using phi29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res 11:1095–1099.

6. Kim K-H, Bae J-W. 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. Appl Environ Microbiol 77:7663–7668.

7. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, Yutin N, Zerbini FM, Kuhn JH. 2020. Global organization and proposed megataxonomy of the virus world. Microbiol Mol Biol Rev 84.

8. Pénzes JJ, Söderlund-Venermo M, Canuti M, Eis-Hübinger AM, Hughes J, Cotmore SF, Harrach B. 2020. Reorganizing the family Parvoviridae: a revised taxonomy independent of the canonical approach based on host association. Arch Virol 1–14.

9. Tijssen P, Penzes JJ, Yu Q, Pham HT, Bergoin M. 2016. Diversity of small, single-stranded DNA viruses of invertebrates and their chaotic evolutionary past. J Invertebr Pathol 140:83–96.

10. Liu H, Fu Y, Xie J, Cheng J, Ghabrial SA, Li G, Peng Y, Yi X, Jiang D. 2011. Widespread endogenization of densoviruses and parvoviruses in animal and human genomes. J Virol JVI–00828.

Journal of Virology

409 11. Hewson I, Button JB, Gudenkauf BM, Miner B, Newton AL, Gaydos JK, Wynne J, Groves CL, Hendler G, Murray M,
410 others. 2014. Densovirus associated with sea-star wasting disease and mass mortality. Proc Natl Acad Sci 111:17278–
411 17283.

412 12. François S, Filloux D, Roumagnac P, Bigot D, Gayral P, Martin DP, Froissart R, Ogliastro M. 2016. Discovery of
413 parvovirus-related sequences in an unexpected broad range of animals. Sci Rep 6.

414 13. Jackson EW, Pepe-Ranney C, Johnson MR, Distel DL, Hewson I. 2020. A highly prevalent and pervasive densovirus
415 discovered among sea stars from the North American Atlantic Coast. Appl Environ Microbiol.

416 14. Gudenkauf BM, Eaglesham JB, Aragundi WM, Hewson I. 2014. Discovery of urchin-associated densoviruses (family
417 Parvoviridae) in coastal waters of the Big Island, Hawaii. J Gen Virol 95:652–658.

418 15. Bucci C, Francoeur M, McGreal J, Smolowitz R, Zazueta-Novoa V, Wessel GM. 2017. Sea Star Wasting Disease in
419 Asterias forbesi along the Atlantic Coast of North America. PLoS ONE 12:e0188523.

420 16. Hewson I, Sullivan B, Jackson EW, Xu Q, Long H, Lin C, Quijano Cardé EM, Seymour J, Siboni N, Jones MRL, Sewell
421 MA. 2019. Perspective: Something Old, Something New? Review of Wasting and Other Mortality in Asteroidea
422 (Echinodermata). Front Mar Sci 6.

17. Hewson I, Bistolas KSI, Quijano Cardé EM, Button JB, Foster PJ, Flanzenbaum JM, Kocian J, Lewis CK. 2018. Investigating the Complex Association Between Viral Ecology, Environment, and Northeast Pacific Sea Star Wasting. Front Mar Sci 5.

18. Miner CM, Burnaford JL, Ambrose RF, Antrim L, Bohlmann H, Blanchette CA, Engle JM, Fradkin SC, Gaddam R, Harley CD. 2018. Large-scale impacts of sea star wasting disease (SSWD) on intertidal sea stars and implications for recovery. PLoS One 13:e0192870.

19. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F. 2009. Laboratory procedures to generate viral metagenomes. Nat Protoc 4:470–483.

20. Bushnell B. 2014. BBTools software package.

21. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, others. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.

22. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, others. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

438 23. Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–
439 3415.

440 24. Gordon A, Hannon GJ. 2010. Fastx-toolkit. FASTQA Short-Reads Preprocessing Tools Unpubl Httphannonlab Cshl
441 Edufastxtoolkit 5.

442 25. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics
443 30:2114–2120.

444 26. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q.
445 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol
446 29:644.

447 27. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59.

448 28. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res
449 32:1792–1797.

450 29. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to
451 estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321.

30. Letunic I, Bork P. 2006. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23:127–128.

31. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28:3150–3152.

32. Cotmore SF, Agbandje-McKenna M, Chiorini JA, Mukha DV, Pintel DJ, Qiu J, Soderlund-Venermo M, Tattersall P, Tijssen P, Gatherer D. 2014. The family parvoviridae. Arch Virol 159:1239–1247.

33. Tijssen P, Bando H, Li Y, Jousset F, Zadori Z, Fediere G, El-Far M, Szelei J, Bergoin M. 2005. Evolution of densoviruses. Parvoviruses 5:55–60.

34. Bergoin M, Tijssen P. 2010. Densoviruses: a Highly Diverse Group. Insect Virol 59.

35. Zolfo M, Pinto F, Asnicar F, Manghi P, Tett A, Bushman FD, Segata N. 2019. Detecting contamination in viromes using ViromeQC. Nat Biotechnol 37:1408–1412.

36. Bowater RO, Wingfield M, Fisk A, Condon KM, Reid A, Prior H, Kulpa EC. 2002. A parvo-like virus in cultured redclaw crayfish Cherax quadricarinatus from Queensland, Australia. Dis Aquat Organ 50:79–86.

37. Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X, Li C-X, Qin X-C, Li J, Cao J-P, Eden J-S. 2016. Redefining the invertebrate RNA virosphere. Nature 540:539–543.

38. Shi M, Zhang Y-Z, Holmes EC. 2018. Meta-transcriptomics and the evolutionary biology of RNA viruses. Virus Res 243:83–90.

39. Waldron FM, Stone GN, Obbard DJ. 2018. Metagenomic sequencing suggests a diversity of RNA interference-like responses to viruses across multicellular eukaryotes. PLoS Genet 14:e1007533.

40. ter Horst AM, Nigg JC, Dekker FM, Falk BW. 2019. Endogenous viral elements are widespread in arthropod genomes and commonly give rise to PIWI-interacting RNAs. J Virol 93:e02124–18.

41. Tattersall P. 2006. The evolution of parvovirus taxonomy. Parvoviruses 1:5–14.

42. Neuwald AF, Aravind L, Spouge JL, Koonin EV. 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res 9:27–43.

43. Gorbalenya AE, Koonin EV, Wolf YI. 1990. A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS Lett 262:145–148.

44. Hickman AB, Dyda F. 2005. Binding and unwinding: SF3 viral helicases. Curr Opin Struct Biol 15:77–85.

45. Herz M, Brehm K. 2019. Evidence for densovirus integrations into tapeworm genomes. Parasit Vectors 12:560.

46. Parker BJ, Brisson JA. 2019. A laterally transferred viral gene modifies aphid wing plasticity. Curr Biol 29:2098–2103.

Figure legends

Figure 1: Reanalysis of metagenomic data presented in Hewson *et al.*, 2014. SSaDV is one of ten densoviruses present in the data set and based on read mapping analysis (≥95% read identity) and is not more abundant in symptomatic compared to asymptomatic individuals. (A) Relative abundance of all reads recruited to densovirus genomes. N.S. (no significance) based on welch two sample t-test (p = 0.7697, df = 25.137, t = -0.29592). (B) Read recruitment separated by densovirus genotype.

Figure 2: SSaDV is broadly distributed across the northeastern Pacific Ocean and putatively has a wide tissue tropism. White dots on map indicated PCR positive sample, and the size of the dot corresponds to total number of PCR positive samples at each site. Tissue tropism assessed from 3 sea star species collected from one site (Langley Harbor, Washington). Colors of each bar corresponds to the anatomical region in the sea star illustration. Prevalence defined as the number of PCR positive samples divided by the total number of samples tested for each tissue.

Figure 3: Sea star-associated densoviruses are genetically diverse and are not monophyletic. (A) Cladogram of a maximum likelihood phylogenetic tree of densoviruses based on alignment of amino acid sequences from NS1 gene. Collapsed nodes represent densovirus genera while all other branches belong to the genus *Ambidensovirus*. Red names indicate genomes

discovered in this study. White circles represent 90-100% bootstrapped support. (B) Representative densovirus genome showing genome organization. (C) Histograms of nucleotide and amino acid pairwise identity comparisons between all sea star-associated densoviruses for NS1, NS3, and VP ORFs. Dotted lines indicate mean pairwise identity.

Figure 4: Sea star-associated densoviruses exhibit two genome organizations and contain motifs typical of densoviruses. Triangles indicate position of amino acid motifs, replication initiation motifs, and Walker box ATPase motifs in densoviruses. Consensus sequences above sequence logos are defined by a 90% identity agreement among all sea star-associated densovirus.

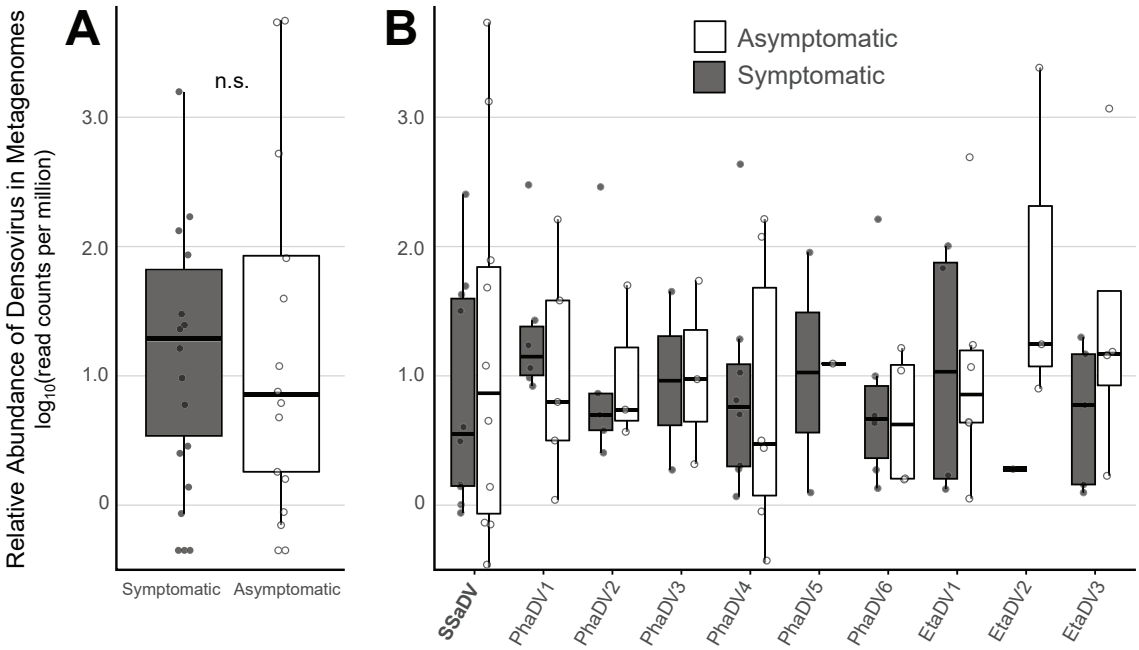
Figure 5: An overview of endogenized densovirus elements (EVE) illustrating the conserved existence of Walker box ATPase motifs. EVEs found in this study are shown in red. Sequences in alignment are grouped by host phylum. Sequences labeled by host species and by origin of sequence from host genome (G), extant virus (V), and host transcriptome (T), and NCBI accession number. Amino acids in bold indicate a sequence identity 75% or greater within the alignment. The blue highlighted region denotes the NTP-binding and helicase region containing Walker A, B, and C motifs found within the NS1/Rep ORF in densoviruses and parvoviruses.

Table 1: Sea star-associated densovirus genome characteristics and metadata

Host (sea star species)	Animal Collection site	State or Providence, Country	Collection Year	Metaviriome	Virus name	Contig Size (nt)	Average Fold Coverage	Orientation	Ambidensovirus Subgroup**	Viral species*	GenBank accession number
Pisaster ochraceus	Santa Cruz	California, USA	2013	RNA	PoaDV1‡	5719	427	Ambisense	A	Asteroid ambidensovirus 2	MT733037
	Olympic National Park	Washington, USA	2013	RNA	PoaDV2‡	5840	63	Ambisense	A	Asteroid ambidensovirus 9	MT733038
	Santa Cruz	California, USA	2013	RNA	PoaDV3‡	5758	134	Ambisense	A	Asteroid ambidensovirus 10	MT733039
	Olympic National Park	Washington, USA	2013	RNA	PoaDV4‡	5469	19	Ambisense	B	Asteroid ambidensovirus 11	MT733040
	Olympic National Park	Washington, USA	2013	RNA	PoaDV5‡	5415	15	Ambisense	A	Asteroid ambidensovirus 12	MT733041
	Olympic National Park	Washington, USA	2013	RNA	PoaDV6‡	5340	27	Ambisense	A	Asteroid ambidensovirus 5	MT733042
	Olympic National Park	Washington, USA	2013	RNA	PoaDV7‡	5052	30	Ambisense	B	Asteroid ambidensovirus 13	MT733043
	Santa Cruz	California, USA	2013	RNA	PoaDV8‡	5584	64	Ambisense	A	Asteroid ambidensovirus 14	MT733044
	Santa Cruz	California, USA	2013	RNA	PoaDV9‡	4827	15	Ambisense	B	Asteroid ambidensovirus 7	MT733045
	Olympic National Park	Washington, USA	2013	RNA	PoaDV10‡	3264	18	Ambisense	NA	Asteroid ambidensovirus 15	MT733046
	Olympic National Park	Washington, USA	2013	RNA	PoaDV11‡	5095	51	Ambisense	B	Asteroid ambidensovirus 8	MT733047
	Santa Cruz	California, USA	2013	RNA	PoaDV12‡	5270	451	Ambisense	A	Asteroid ambidensovirus 6	MT733048
	Santa Cruz	California, USA	2018	RNA	PoaDV13	5963	30	Ambisense	A	Asteroid ambidensovirus 16	MT733049
	Santa Cruz	California, USA	2018	RNA	PoaDV14	5274	72	Ambisense	B	Asteroid ambidensovirus 17	MT733050
Leptasteria sp.	Palmer's Point & Pigeon Point	California, USA	2017	RNA	LhaDV1	3956	35	Ambisense	NA	Asteroid ambidensovirus 7	MT733022
Pycnopodia helianthoides	Seattle Aquarium	Washington, USA	2013	RNA	PhaDV1†	5665	47	Ambisense	A	Asteroid ambidensovirus 2	MT733031
	Seattle Aquarium	Washington, USA	2013	RNA	PhaDV2†	5326	27	Ambisense	A	Asteroid ambidensovirus 5	MT733032
	Seattle Aquarium	Washington, USA	2013	RNA	PhaDV3†	4168	15	Ambisense	NA	Asteroid ambidensovirus 18	MT733033
	Seattle Aquarium	Washington, USA	2013	RNA	SSaDV†	5663	34	Ambisense	A	Asteroid ambidensovirus 1	
	Seattle Aquarium	Washington, USA	2013	RNA	PhaDV4†	5485	58	Ambisense	A	Asteroid ambidensovirus 6	MT733034
	Seattle Aquarium	Washington, USA	2013	RNA	PhaDV5†	3391	16	NA	A	Only VP gene	MT733035
	Seattle Aquarium	Washington, USA	2013	RNA	PhaDV6†	3446	27	Ambisense	NA	Asteroid ambidensovirus 8	MT733036
	Burrard Inlet	British Columbia, Canada	2013	DNA	SSaDV†	6053	246	Ambisense	A	Asteroid ambidensovirus 1	MT733051
Evasterias troschelii	Cape Roger Curtis	British Columbia, Canada	2013	DNA	SSaDV†	5206	27	Ambisense	A	Asteroid ambidensovirus 1	

	Cape Roger Curtis	British Columbia, Canada	2013	DNA	EtaDV1†	5601	15	Ambisense	B	Asteroid ambidensovirus 19	MT733014
	Cape Roger Curtis	British Columbia, Canada	2013	DNA	EtaDV2†	5700	38	Ambisense	B	Asteroid ambidensovirus 20	MT733015
	Cape Roger Curtis	British Columbia, Canada	2013	DNA	EtaDV3†	5460	22	Ambisense	B	Asteroid ambidensovirus 21	MT733016
Neosmilaster georgianus	Palmer Station	Antarctica	2017	RNA	NgaDV1	5605	122	Ambisense	B	Asteroid ambidensovirus 3	MT733025
	Palmer Station	Antarctica	2017	RNA	NgaDV2	5383	395	Ambisense	B	Asteroid ambidensovirus 3	MT733026
	Palmer Station	Antarctica	2017	RNA	NgaDV3	5328	388	Ambisense	B	Asteroid ambidensovirus 22	MT733027
	Palmer Station	Antarctica	2017	RNA	NgaDV4	5352	36423	Ambisense	A	Asteroid ambidensovirus 23	MT733028
	Palmer Station	Antarctica	2017	RNA	NgaDV5	4832	15	Ambisense	B	Asteroid ambidensovirus 4	MT733029
	Palmer Station	Antarctica	2017	RNA	NgaDV6	4886	63	Ambisense	A	Asteroid ambidensovirus 24	MT733030
Labidiaster annulatus	Palmer Station	Antarctica	2017	RNA	LaaDV1	5872	443	Ambisense	B	Asteroid ambidensovirus 25	MT733017
	Palmer Station	Antarctica	2017	RNA	LaaDV2	5240	214	Ambisense	B	Asteroid ambidensovirus 4	MT733018
	Palmer Station	Antarctica	2017	RNA	LaaDV3	5018	501	Ambisense	B	Asteroid ambidensovirus 26	MT733019
	Palmer Station	Antarctica	2017	RNA	LaaDV4	4972	68	Ambisense	B	Asteroid ambidensovirus 27	MT733020
	Palmer Station	Antarctica	2017	RNA	LaaDV5	5413	263	Ambisense	B	Asteroid ambidensovirus 3	MT733021
Luidia maculata	Hong Kong	China	2014	DNA	LmaDV1‡	5446	16	Ambisense	B	Asteroid ambidensovirus 28	MT733023
Astropecten polyacanthus	Hong Kong	China	2014	DNA	ApaDV1‡	3061	8	Ambisense	NA	Asteroid ambidensovirus 29	MT733013
Mediaster aequalis	Ketchikan	Alaska, USA	2016	RNA	MaaDV1	5956	1220	Ambisense	A	Asteroid ambidensovirus 2	MT733024
Asterias forbesi	Nahant	Massachusetts, USA	2015	DNA	AfaDV	6089	454	Ambisense	A	Asteroid ambidensovirus 1	MN190158

- * Viral species defined by a pairwise amino acid sequence identify of NS1 >85% (32).
- ** Ambidensovirus subgroups defined in (33)
- † Viral genotypes discovered from viral metagenomes prepared in (11)
- ‡ Viral genotypes discovered from viral metagenomes prepared in (16)



JVI Journal of Virology

Journal of Virology