A highly prevalent and pervasive densovirus discovered among sea stars from the North American Atlantic Coast

Elliot W. Jackson¹*, Charles Pepe-Ranney², Mitchell R. Johnson¹, Daniel L. Distel³, Ian Hewson¹

¹ Department of Microbiology, Cornell University, Ithaca NY 14853, USA
² AgBiome Inc., 104 T.W. Alexander Drive, Research Triangle Park, NC 27709
³ Ocean Genome Legacy Center of New England BioLabs, Northeastern University Marine Science Center, Nahant MA, 01908

* Correspondence: ewj34@cornell.edu; Tel.: +1-231-838-6042
Abstract:
The etiology of Sea Star Wasting Syndrome is hypothesized to be caused by a densovirus, SSaDV, that has previously been reported on the Pacific and Atlantic coasts of the United States. In this study, we reevaluated the presence of SSaDV among sea stars from the North American Atlantic Coast and in doing so discovered a novel densovirus we have named Asterias forbesi associated densovirus (AfaDV) that shares 78% nucleotide pairwise identity to SSaDV. In contrast to previous studies, SSaDV was not detected in sea stars from the North American Atlantic Coast. Using a variety of PCR-based techniques, we investigated the tissue tropism, host specificity, and prevalence of AfaDV among populations of sea stars at five locations along the Atlantic Coast. AfaDV was detected in three sea star species (Asterias forbesi, Asterias rubens, and Henricia sp.) found in this region and was highly prevalent (>80% of individuals tested, n=134), among sampled populations. AfaDV was detected in the body wall, gonads, and pyloric caeca (digestive gland) of specimens but was not detected in their coelomic fluid. A significant difference in viral load was found between tissue types with the pyloric caeca having the highest viral loads. Further investigation of Asterias forbesi gonad tissue found germline cells (oocytes) to be virus positive, suggesting a potential route of vertical transmission. Taken together, these observations show that the presence of AfaDV is not an indicator of Sea Star Wasting Syndrome because AfaDV is a common constituent of these animals’ microbiome, regardless of health.

Importance:
Sea Star Wasting Syndrome is a disease primarily observed on the Pacific and Atlantic coast of North America that has significantly impacted sea star populations. The etiology of this disease is unknown though hypothesized to be caused by a densovirus, SSaDV. However, previous
studies have not found a correlation between SSaDV to Sea Star Wasting Syndrome on the North American Atlantic Coast. This study suggests that this observation may be explained by the presence of a genetically similar densovirus, AfaDV, that may have confounded previous studies. SSaDV was not present in sea stars screened in this study, and instead AfaDV was commonly found in sea star populations across the New England region with no apparent signs of disease. These results suggest that sea star densovirus may be common constituents of the animal’s microbiome, and the diversity and extent of these viruses among wild populations may be greater than previously recognized.

Introduction:

Densovirus, also known as densonucleosis viruses, are icosahedral, non-enveloped viruses that have monopartite linear single-stranded DNA genomes that are typically 4-6 kb packaged in a 20-25 nm diameter capsid shell (1, 2). Densovirus belong to the subfamily Densovirinace which is part of the Parvoviridae family and are known to infect arthropods, specifically insects and shrimp (1). Prior to the advent of high throughput sequencing technology, the discovery of densovirus was driven by investigations of epizootics occurring in laboratory populations and breeding facilities of economically important invertebrates (e.g. silkworms, crickets, shrimp) or through infected cell lines (e.g. mosquito C6/36 cell line) (3–7). The majority of densovirus isolated to date share a common pathology, causing hypertrophied nuclei in affected tissues, and are generally more virulent at early life stages of its host (1). Densovirus have also been shown to be mutualists. For example, sublethal infections in rosy apple aphids are correlated with a winged phenotype that has a lower fecundity compared to non-winged aphids (8). The cost of infection lowers the fecundity of the individual to promote the
growth of wings that increases mobility and the potential for the host and the virus to disperse 
(8). Although densoviruses have been primarily studied in insects and crustaceans, analysis of 
transcriptomic datasets and viral metagenomes prepared from metazoan tissues have found 
endogenous and exogenous densovirus sequences from a much wider host range (9–13). These 
findings suggest that densoviruses may be common constituents of many invertebrate viromes. 

Densovirus sequences have recently been recovered from echinoderm tissues and have 
been implicated as potential pathogens though their relationship to echinoderms is unknown (10, 
11, 14). In 2013-2015 a mortality event termed Sea Star Wasting Disease or Syndrome (SSWS) 
(also referred to as Asteroid Idiopathic Wasting Syndrome) affected >20 sea star species along a 
broad geographic range from California to Alaska. A densovirus (Sea Star associated Densovirus 
or SSaDV) was discovered and hypothesized to cause SSWS (11, 14). Previous work has 
reported the presence of SSaDV, using PCR and qPCR (11, 15, 16), among sea stars on the 
North American Atlantic Coast though was not found to be significantly correlated with SSWS 
(15, 16). The cause of SSWS among echinoderms on the Atlantic Coast is however hypothesized 
to be viral in nature (16).

We sought to reinvestigate the presence of SSaDV and more generally survey the viral 
diversity of densoviruses in sea stars inhabiting the North American Atlantic Coast. Viral 
metagenomes were prepared from Asterias forbesi, a common sea star found in the sub-tidal 
environment from the North American Atlantic Coast which led to the discovery of a novel sea 
star densovirus hereafter referred to as Asterias forbesi associated densovirus or AfaDV. AfaDV 
is the second sea star associated densovirus to be discovered in sea stars thus far. In contrast to 
previous work, we did not find any evidence of SSaDV among sea stars on the Atlantic Coast 
through metagenomic analysis and PCR surveys (11, 15, 16). Using a variety of PCR
techniques, we investigated the geographic distribution of AfaDV, the host specificity, tissue
tropism, and the potential for vertical transmission. Our results show that AfaDV has a broad
geographic range, is not species specific, has a wide tissue tropism, and is potentially vertically
transmitted.

Methods:

Viral metagenomic preparation and bioinformatic analysis:

Viral metagenomes were prepared using a protocol that was adapted and modified from
existing laboratory protocols (17). The viral metagenomes used in this study have been
previously analyzed and reported for the presence of circular ssDNA viruses (18). Six Asterias
forbesi that displayed signs characteristic of SSWS (arm detachment, mucoid appearance on
aboral surface, disintegration of epidermal tissue) were collected from Canoe Beach, Nahant
Bay, Massachusetts, USA (42.420889, -70.906416) in September/October 2015 (19). Animals
were flash frozen in liquid nitrogen upon collection then stored at -80°C until dissection. The
pyloric caeca and body wall tissue from these animals were pooled separately in 0.02µm-
filtered 1X PBS then homogenized in a bleached-cleaned NutriBullet for 60s. Tissue homogenates were
pelleted by centrifugation at 3,000 x g for 180s, and the supernatant was syringe filtered through
Millipore Sterivex-GP 0.22µm polyethersulfone filters into bleach-treated and autoclaved
Nalgene Oak Ridge High-Speed Centrifugation Tubes. Filtered homogenates were adjusted to a
volume of 35mL by adding 0.02µm-filtered 1X PBS and amended with 10% (wt/vol) PEG-8000
and precipitated for 20 hours at 4°C. Insoluble material was then pelleted by centrifugation at
15,000 x g for 30 minutes. The supernatant was decanted and pellets were resuspended in 1mL
of 0.02µm-filtered nuclease-free H2O. The samples were treated with 0.2 volumes (200µl) of
CHCl3, 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 DNA was extracted from 500-μl subsamples using the Zymo Research Viral DNA™ kit following the manufacturer’s protocol and subsequently amplified isothermally at 30°C using Genomiphi Whole Genome Amplification Kit (GE Healthcare, Little Chalfont, UK). Samples were cleaned and concentrated using a Zymo Research DNA Clean and Concentrator™ kit, and DNA was quantified using Quant-iT™ PicoGreen® dsDNA Assay kit (Invitrogen). Samples were prepared for Illumina sequencing using the Nextera XT DNA Library Preparation kit (Illumina, San Diego, CA, USA) prior to 2x250bp paired-end Illumina MiSeq sequencing at Cornell University Core Laboratories Center (Ithaca, NY, USA).

Libraries generated from both samples (pyloric caeca and body wall) were first interleaved into one file. Reads were then trimmed for read quality and Illumina adapters and filtered for phiX contamination then 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 done using the BBtools suite (20). The merged and unmerged read normalized libraries were used for de novo assembly using SPAdes (21). Contigs less than 500nt were discarded after assembly and the remaining contigs were subject to tBLASTx against a curated in-house database of ssDNA viruses (22). Contigs with significant (e-value < 1x10⁻⁸) sequence similarity to SSaDV (“Sea Star associated Densovirus”) were isolated and reads were mapped back to contigs with a minimum identity of 0.95 to obtain average coverage and coverage distribution across the contigs (Figure 1). ORFs
were defined in Geneious with a minimum size of 550nt, and the hairpin structures in the
inverted terminal repeats (ITRs) were determined using Mfold (23, 24). AfaDV genome
sequence has been deposited in GenBank under the accession number MN190158. Metagenomic
libraries have been deposited in GenBank under BioProject ID PRJNA555067. All statistical and
bioinformatics analyses can be found at https://github.com/ewj34/AfDV-Viral-Metagenome.

Phylogenetic relationships among AfaDV and 45 densovirus genomes were inferred by a
maximum likelihood method, implementing SMS (smart model selection) in PhyML 3.0 using a
MUSCLE amino acid sequence alignment of NS1 with default parameters (25, 26). The region
of NS1 used for alignment (sequence length 437.7 ± 50; mean ± SD) spanned Motif I of the RC
endonuclease domain to Motif C of the SF3 helicase domain. Branch support was determined by
bootstrapping at 100 iterations. The phylogenetic tree was visualized and edited using iTOL (27).

Specimen collection, nucleic acid extraction, cDNA synthesis:

Three species of sea stars - Asterias forbesi, Asterias rubens, and Henricia sp. - were
collected from five locations along the Atlantic Coast of the United States from 2012 to 2019
(Supplementary Table 1) (19). 83 Asterias forbesi, 16 Henricia sp, and 35 Asterias rubens were
collected totaling 134 sea stars. Prior to vivisection or dissection, animal length was measured by
total diameter (i.e. ray to ray). Animals were either vivisected immediately after collection and
flash frozen in liquid nitrogen upon collection then stored at -80°C until dissection or
cryopreserved and stored at -20°C until dissection. Coelomic fluid was extracted from only
animals vivisected using a 25G x 1½ (0.5mm x 25mm) needle attached to a 3 mL syringe
inserted through the body wall into the coelomic cavity. Gonads, body wall, and pyloric caeca
were collected from all animals vivisected, but not every tissue type was collected from animals
that were dissected (Supplementary Table 1). In total 368 samples were collected. DNA was
extracted from tissues (14 - 200 mg wet weight) and coelomic fluid (140 μl - 1000 μl) samples using Zymo Research Tissue & Insect DNA™ kits or Zymo Research Quick DNA Miniprep Plus™ kit following the manufacture’s protocol (Supplementary Table 1). All DNA samples used for qPCR were extracted with the Zymo Research Tissue & Insect DNA™ kit.

Coelomocytes in the coelomic fluid were pelleted by centrifugation at 10,000 x g for 5 minutes then resuspended in 200 μl of 0.02 μm-filtered nuclease-free H₂O prior to DNA extraction. DNA was quantified using Quant-iT™ PicoGreen® dsDNA Assay kit (Invitrogen). RNA was extracted from pyloric caeca, gonads, and body wall samples from Asterias forbesi collected from Nahant, Massachusetts using the Zymo Research Tissue & Insect RNA MicroPrep™ kit. An in-column DNase I digestion was performed following the manufacture’s protocol. The Maxima® First Strand cDNA Synthesis Kit (Thermo Scientific) was used for cDNA synthesis following RNA extraction. Four microliters of eluted RNA from each sample was stored at -80°C and was used as a no reverse transcriptase (no-RT) control for RT-PCR analysis.

Genome verification:

AfaDV was amplified by PCR with overlapping primers to produce 16 amplicons. Amplicons generated from PCR were gel visualized with ethidium bromide, and the remaining PCR product was cleaned and concentrated using a Zymo Research DNA Clean and Concentrator™ kit and submitted for Sanger sequencing at Cornell University Core Laboratories Center (Ithaca, NY, USA). The resulting amplicons were assembled to form a contig spanning nucleotide positions 214 to 5860 (92.7% of total genome length). The assembled contig was identical to the contig generated from the de novo assembly. This process was repeated on tissue samples from the three sea star species collected to verify the viral genome sequence among each
Putative VP and NS genes of AfaDV and SSaDV were inserted into the pGEM-t-Easy vector and used to transform NEB 5-alpha competent cells. Primers, PCR conditions, and plasmid constructs can be found in Table 1/Supplemental Figure 1. PCR amplicons were gel purified, poly(A) tailed, and ligated to pGEM-t-Easy vectors. The pUC19-SSaDV construct was synthesized by Genscript. Plasmid constructs were verified by Sanger sequencing.

Viral prevalence and viral load:
Viral prevalence and viral load were determined using quantitative(q)PCR using TaqMan chemistry. qPCR primers and probe were designed using Primer3 targeting the VP region positions 5634 to 5728 (28). Reaction conditions and primer/probe/oligonucleotide standard sequences can be found in Table 1. All qPCR reactions, including no-template controls, were performed in duplicate on a StepOnePlus™ Real-Time PCR system (Applied Biosystems). A synthetic oligonucleotide sequence spanning the qPCR primer/probe region was used to generate the standard curve using duplicate eight-fold serial dilutions which were included in all qPCR runs. 25µl reaction volumes contained 0.02µl (200 pmol) of each primer and probe, 12.5µl of 2X SsoAdvanced™ Universal Probes Supermix (Bio-Rad), 10.44µl of nuclease-free H₂O, and 2µl of template. Viral quantity was calculated by StepOnePlus™ software (version 2.3 Applied Biosystems) by averaging the cycle threshold between duplicates and interpolating values against the standard curve. Viral quantities were adjusted by extraction volume and standardized by sample weight. Samples were considered positive when technical replicates were both positive and Ct standard deviation was <1.0. The lower limit of detection used in this study was...
Fifteen *Asterias forbesi* were collected from Woods Hole, Massachusetts in June/July 2019. Gonadal tissue and pyloric caeca were taken through a small incision to the arm and manually extracted with forceps (Supplemental Table 1). Isolation of oocytes was performed according to Wessel *et al.*, 2010 (29). Oocytes were isolated by mincing gonadal tissue in 0.2μm filtered sea water then poured through cheesecloth, pelleted by centrifugation, decanted, and pipetted for DNA extraction. DNA extractions were performed using Zymo Research Quick DNA MiniPrep Plus™ kit. PCR reactions included a kit negative control (i.e. extraction blank) and PCR reagent negative control to account for false-positives. PCR cycle conditions can be found in Table 1.

Thirty pyloric caeca samples were screened for SSaDV to determine the presence of SSaDV among North Atlantic sea stars. Ten pyloric caeca samples from Woods Hole, Shoals Marine Lab, and Nahant were chosen. The specificity of primers was validated (i.e. no cross amplification observed) by screening against the appropriate plasmid constructs prior to screening DNA extracts (Supplemental Figure 1). PCR cycle conditions can be found in Table 1.

**Results:**

**Genome analysis and phylogeny:**

High-throughput sequencing of metaviromes prepared from the pyloric caeca and body wall of sea star samples generated $2.42 \times 10^7$ and $7.66 \times 10^6$ reads/library respectively totaling
3.18 x 10^7 reads. SPAdes assembly and annotation of contigs to the curated ssDNA database resulted in one contig that was 6,089nt that was significantly (e-value < 1x10^-8) similar to SSaDV (Figure 1). Read mapping to this contig recruited 16,737 reads that gave an average base coverage of 532 that proportionately made up 0.052% of the total reads (Figure 1). A total of 16,732 reads mapped came from the metavirome prepared from the pyloric caeca and 5 reads mapped from the metavirome prepared from body wall tissue. The contig contained 4 ORFs that putatively encode non-structural proteins (NS1, NS2, and NS3) and a structural protein (VP) (Figure 1). As a whole, the contig encodes all components of the NS and VP cassettes that are characteristic of the genus Ambidensovirus within the subfamily Densovirinae. Phylogenetic analyses (Maximum Likelihood) of the aforementioned sequence indicate that this novel densovirus falls within a well-supported clade that includes other Ambidensoviruses and shares a most recent common ancestry with the previously described sea star densovirus, SSaDV (Figure 2, Supplemental Figure 2). AfaDV and SSaDV share a 77.9% pairwise nucleotide identity across entire genomes though the putative NS1, NS2, and NS3 genes have pairwise nucleotide identities of 88.7%, 88.1%, and 74.4% respectively while the putative VP genes share a 77.1% pairwise nucleotide identity. According to Cotmore et al., paroviruses of the same species encode NS1 proteins with >85% pairwise amino acid sequence identity (30). The pairwise amino acid sequence identity of NS1 between AfaDV and SSaDV is 86.5% which makes AfaDV a new isolate of the same densovirus species. It is likely that the contig is a complete genome because of the presence of the hairpins located at the ends of the inverted terminal repeats (ITRs) which are characteristic of paroviruses (31). However, it is possible the ITRs were not fully completed in silico based off the observation that ambidensovirus ITR nucleotide lengths are typically >500nt which raises the possibility that this genome is not complete (1). Nevertheless, the length
of the ITRs are 271nt long, and the terminal nucleotides of the ITRs form canonical hairpin structures that are 93nt long which are thermodynamically favorable ($\Delta G = -52.71$) (Figure 1).

Tissue tropism, prevalence, and biogeography:

SSaDV was not detected in any of the pyloric caeca samples screened. AfaDV was detected from sea stars collected from 2012 – 2019 from four of the five locations (Appledore Island, Maine, Nahant, Massachusetts, Woods Hole, Massachusetts, and the Mystic Aquarium, Connecticut) (Figure 3). AfaDV was not detected from 6 Asterias forbesi collected from Bar Harbor, ME (Figure 3). The overlapping PCR primer set successfully amplified AfaDV sequences from the three sea star species collected in this study. To assess tissue tropism, we screened pyloric caeca, gonads, body wall, and coelomic fluid from animals via qPCR and RT-PCR. AfaDV was detected more frequently in the pyloric caeca (86%) > body wall (70%) > gonads (57%) and was not detected in any coelomic fluid samples (Figure 4). Viral load was significantly different among tissue types (ANOVA, $F_{2,197} = 29.11$, $p = 8.38 \times 10^{-12}$; Figure 4).

AfaDV log$_{10}$ transformed values were significantly greater in the pyloric caeca (3.3 ± 0.94; mean ± SD) than in the body wall (2.6 ± 0.66, Games-Howell, $p = 5.6 \times 10^{-8}$) and gonads (2.3 ± 0.52, Games-Howell, $p = 5.2 \times 10^{-11}$). AfaDV load was not significantly different between gonads and body wall (Games-Howell, $p = 0.13$). Viral load was also significantly correlated with animal length (Pearson’s correlation, $r = 1.256 \times 10^{-7}$) and had a negative association with animal length ($r = -0.383$) (Figure 4). Average DNA concentration differences found between tissue types did not reflect the trends found in viral loads across tissue types (Supplemental Figure 3). RT-PCR was performed on RNA extracted from pyloric caeca, gonads, and body wall samples to determine transcription of AfaDV in these tissues. Viral transcripts were detected in all three tissue types (Supplemental Figure 4).
Vertical transmission of AfaDV:

The pyloric caeca of 15 Asterias forebdi from Woods Hole, Massachusetts were virus positive via PCR. 5/10 oocytes isolated from females were virus positive and 2/5 gonadal tissue from male sea stars were virus positive (Supplemental Figure 5).

Discussion:
The discovery of SSaDV and detection in sea stars from the Northwest Pacific and Northwest Atlantic suggest that SSaDV is associated with SSWS in disparate geographic regions (11). Subsequent investigations (15, 16) have supported an association between the occurrence of SSaDV and the incidence of SSWS among sea stars in the Atlantic using primers that were presumably specific to SSaDV. Previous documentation of SSaDV in Atlantic sea stars however may be confounded by spurious amplification and/or the presence of a genetically similar densovirus. Indeed, the nucleotide similarity of AfaDV and SSaDV suggests that previous primers may have been insufficient in distinguishing these two genotypes which led to the conclusion that SSaDV is associated with sea stars on the Atlantic Coast. By validating the specificity of our primers, we tested the presence of both genotypes and did not find evidence for the presence of SSaDV in sea stars on the Atlantic Coast. These results suggest that SSaDV is limited to sea stars in the Northwest Pacific which implies any correlation to SSWS outside of this region is unlikely. Although SSWS is broadly defined and is most notably observed in the Northwest Pacific, it has been observed in the Shetland Islands near Antarctica, northern Australia, and the Yellow Sea, China (32). If densoviruses are correlated with SSWS in disparate geographic regions, unique densovirus genotype(s) may also exist in those regions. Further efforts to document the diversity and biogeography of these viruses may help to elucidate these correlations.
Currently, it is unclear what environmental or host-specific factors shape the biogeography of AfaDV and SSaDV, but the diversity and prevalence of these viruses among wild populations may be underappreciated. Similar to SSaDV, AfaDV is not associated with one species and can be found across a large geographic range (Figure 3). To accurately document the prevalence of AfaDV, we first investigated tissue tropism, which has not been established for this virus-host system, to identify the best tissue type for viral detection (Figure 4). Densoviruses typically have a wide tissue tropism in arthropods and can actively replicate in most tissues though replication can be exclusively limited to certain tissues depending on the viral genotype.

For example, densoviruses in the genus Iteradensovirus replicate exclusively and/or predominantly in midgut epithelium cells of their hosts while G. mellonella and J. coenia ambidensoviruses replicate in almost all tissues except the midgut epithelium (1, 33). Using qPCR, PCR, and RT-PCR, we detected AfaDV in the pyloric caeca, gonads, and body wall with varying degrees of prevalence but no positive detection was found in the coelomic fluid (Figure 4, Supplemental Figure 4). Comparative analysis across tissue types showed significantly higher viral loads per unit sample weight in the pyloric caeca than other tissue types (Figure 4). This difference was not found to be a reflection in DNA concentrations between tissue types which suggests the viral loads across tissue types is due to a biological difference (Supplemental Figure 3). It should be noted that our investigation of tissue tropism did not include stomach, intestinal caeca, and radial nerves. Analysis of these tissues would be needed to establish complete tissue tropism in addition to using microscopy techniques (in-situ hybridization or immunohistochemistry) to further validate these results. Because viral transcripts were detected in the pyloric caeca, gonads, and body wall, the differences in viral load might result from the susceptibility of cell types, rates of cellular division among tissues or accessibility of tissue to the
host immune system. Given that parvoviruses replicate during the S phase of the cell cycle, the
trends in viral load across tissues likely reflect differences in cellular proliferation between these
tissues (31). We hypothesize that the pyloric caeca has a larger proportion of dividing cells
relative to other tissues, thereby explaining the observed differences in viral load. Similarly, the
correlation between viral load and animal length could be a reflection of a greater proportion of
 cellular division in growing individuals.

The detection of AfaDV in the same populations over a two to three-year time span
suggest that sea stars can maintain persistent densovirus infections. Persistent infections are
common among vertebrate paroviruses and densoviruses (34–37). Paroviruses replicate
passively and are generally more virulent in fetal and juvenile organisms while adults can
maintain persistent infections without showing any clinical signs. For example, shrimp
densoviruses can cause acute infections in juveniles but individuals that survive can carry the
virus for life transmitting it vertically and horizontally (38). Infected adults rarely show signs of
disease even in individuals with heavy infections (37). Persistent infections have been reported in
other virus-host invertebrate systems. Notably viruses that infect Apis mellifera, the European
honey bee, form persistent infections that can lead to acute infections under certain conditions
(39, 40). The persistence of AfaDV among host populations brings into question the capacity of
sea star densoviruses to cause acute infections under particular circumstances. Such
circumstances might include one or a combination of biotic and abiotic factors such as the
 nutritional status of the host, temperature fluctuations, microbial dysbiosis, or the presence of a
vector (e.g. Varroa destructor). Considering the high prevalence of AfaDV or SSaDV among
sea stars with no obvious signs of infection, the association to SSWS, if any, may be the result of
a combined effect of viral infection(s) and biotic or abiotic changes.
The presence of AfaDV in DNA extracted from sea star oocytes suggest that AfaDV can infect the animal’s germline cells (Supplemental Figure 5). Active or latent infection during embryonic development, larval growth, and metamorphosis could facilitate vertical transmission, an efficient mechanism to achieve high prevalence in adults. Further experimental and microscopic evidence will be necessary to fully establish this route of transmission. The observation of AfaDV in oocytes also suggests these cells may be permissive. Because the whole genome of AfaDV was recovered in silico, a synthetic clone could be constructed and microinjected into developing embryos to test the permissiveness of these cells. Given that echinoderms are model organisms in developmental biology and infectious clones are routinely used to study densovirus biology, the tractability of this approach is promising (41–45, 29).

These experiments are currently being tested using a SSaDV clone as part of an ongoing investigation to further understand the relationship of densoviruses to sea stars.

Here we report the discovery of a novel sea star densovirus from Asterias forbesi. Phylogenetic analysis demonstrates that this virus is closely related to the previously discovered sea star densovirus, SSaDV. Our investigation did not find evidence for the presence of SSaDV in specimens from the North Atlantic, suggesting that SSaDV is limited to Northwest Pacific sea stars. PCR based approaches were used to investigate tissue tropism, prevalence among healthy sea star populations in the environment, and potential vertical transmission of AfaDV. We found AfaDV to have a broad geographic range that spreads across Connecticut to Maine and is found in high prevalence among populations. The vertical transmission of AfaDV may explain these high prevalence rates. The results of this study further our understanding of the association between densoviruses and echinoderms beyond the context of disease. The prevalence and pervasiveness of AfaDV among wild populations suggests that these viruses might form
commensal or mutualistic relationships with their hosts. The pathogenicity and interaction of AfaDV at the cellular, larval and adult stage cannot be inferred from these data alone, but it appears that densoviruses may be common constituents of these animal’s microbiome.

**Funding:**

This work was supported by NSF grants OCE-1537111 and OCE-1737127 awarded to IH. This work was also supported by the Cornell Atkinson Center’s Sustainable Biodiversity Fund and Andrew W. Mellon Student Research Grant awarded to EWJ.

**Acknowledgments:**

The authors thank Dr. Holly Lutz, Taylor Dueweke, Dr. Lisa Abbo, Dr. Charlotte Seid, Dr. Kathy Tuxbury, Kathryn Bland, and Dr. James Coyer for their assistance in animal collection. The staff at Ocean Genome Legacy, Shoals Marine Laboratory, and the Marine Biological Laboratory at Woods Hole. The authors also thank Dr. Kalia Bistolas, Dr. Gary Wessel, and Dr. Nathalie Oulhen for their assistance and suggestions with laboratory work and manuscript comments.

**Literature Cited**

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

2. Carlson J, Suchman E, Buchatsky L. 2006. Densoviruses for control and genetic manipulation of mosquitoes. Adv Virus Res 68:361–392.

3. Meynadier G, Vago C, Plantevin G, Atger P. 1964. Virose d’un type inhabituel chez le lépidoptère Galleria mellonella L. Rev Zool Agric Appl 63:207–208.
4. Lightner DV. 1996. Epizootiology, distribution and the impact on international trade of two penaeid shrimp viruses in the Americas. Rev Sci Tech Int Off Epizoot 15:579–601.

5. Liu K, Li Y, Jousset F-X, Zadori Z, Szélei J, Yu Q, Pham HT, Lépine F, Bergoin M, Tijssen P. 2011. The Acheta domesticus Densovirus, Isolated from the European House Cricket, Has Evolved an Expression Strategy Unique among Paroviruses. J Virol 85:10069–10078.

6. Jousset F-X, Baquerizo E, Bergoin M. 2000. A new densovirus isolated from the mosquito Culex pipiens (Diptera: Culicidae). Virus Res 67:11–16.

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

8. Ryabov EV, Keane G, Naish N, Evered C, Winstanley D. 2009. Densovirus induces winged morphs in asexual clones of the rosy apple aphid, Dysaphis plantaginea. Proc Natl Acad Sci 106:8465–8470.

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

10. Gudenkauf BM, Eaglesham JB, Aragundi WM, Hewson I. 2014. Discovery of urchin-associated densoviruses (family Parvoviridae) in coastal waters of the Big Island, Hawaii. J Gen Virol 95:652–658.
11. Hewson I, Button JB, Gudenkauf BM, Miner B, Newton AL, Gaydos JK, Wynne J, Groves CL, Hendler G, Murray M, others. 2014. Densovirus associated with sea-star wasting disease and mass mortality. Proc Natl Acad Sci 111:17278–17283.

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

13. Kang Y-J, Huang W, Zhao A-li, Lai D-D, Shao L, Shen Y-Q, Deng X, Delwart E, Zhang W. 2017. Densoviruses in oyster Crassostrea ariakensis. Arch Virol 162:2153–2157.

14. 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.

15. Fahsbender E, Hewson I, Rosario K, Tuttle AD, Varsani A, Breitbart M. 2015. Discovery of a novel circular DNA virus in the Forbes sea star, Asterias forbesi. Arch Virol 160:2349–2351.

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

17. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F. 2009. Laboratory procedures to generate viral metagenomes. Nat Protoc 4:470–483.
18. Jackson EW, Bistolas KS, Button JB, Hewson I. 2016. Novel Circular Single-Stranded DNA Viruses among an Asteroid, Echinoid and Holothurian (Phylum: Echinodermata). PloS One 11:e0166093.

19. OGL Accession ID number [S22798-S22803,S23233-S23273. Ocean Genome Legacy Database, The Ocean Genome Legacy Center of New England Biolabs, Northeastern University. Published on the web at: https://www.northeastern.edu/ogl/catalog.

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. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

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

24. 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.

25. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797.
26. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321.

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

28. Rozen S, Skaletsky H. 1999. Primer3 on the WWW for general users and for biologist programmers. Bioinformatics Methods Protoc 365–386.

29. Wessel GM, Reich AM, Klatsky PC. 2010. Use of Sea Stars to Study Basic Reproductive Processes. Syst Biol Reprod Med 56:236–245.

30. 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.

31. Cotmore S, Tattersall P. 2005. A rolling-hairpin strategy: basic mechanisms of DNA replication in the parvoviruses, p. 171–181. In Paroviruses. London: Hoddler Arond. Oxford University Press Inc.

32. Hewson I, Sullivan B, Jackson EW, Xu Q, Long H, Lin C, Quijano Cardé EM, Seymour J, Siboni N, Jones MRL, Sewell MA. 2019. Perspective: Something Old, Something New? Review of Wasting and Other Mortality in Asteroidea (Echinodermata). Front Mar Sci 6.
33. Mutuel D, Ravallec M, Chabi B, Multeau C, Salmon J-M, Fournier P, Ogliastro M. 2010. Pathogenesis of Junonia coenia densovirus in Spodoptera frugiperda: A route of infection that leads to hypoxia. Virology 403:137–144.

34. Jacoby RO, Johnson EA, Paturzo FX, Gaertner DJ, Brandsma JL, Smith AL. 1991. Persistent rat parvovirus infection in individually housed rats. Arch Virol 117:193–205.

35. Bloom ME, Kanno H, Mori S, Wolfinbarger JB. 1994. Aleutian mink disease: puzzles and paradigms. Infect Agents Dis 3:279–301.

36. Sandals WC, Povey RC, Meek AH. 1995. Prevalence of bovine parvovirus infection in Ontario dairy cattle. Can J Vet Res 59:81.

37. Flegel TW. 1997. Major viral diseases of the black tiger prawn (Penaeus monodon) in Thailand. World J Microbiol Biotechnol 13:433–442.

38. Flegel TW. 2007. Update on viral accommodation, a model for host-viral interaction in shrimp and other arthropods. Dev Comp Immunol 31:217–231.

39. Tentcheva D, Gauthier L, Zappulla N, Dainat B, Cousserans F, Colin ME, Bergoin M. 2004. Prevalence and seasonal variations of six bee viruses in Apis mellifera L. and Varroa destructor mite populations in France. Appl Environ Microbiol 70:7185–7191.

40. Prisco GD, Zhang X, Pennacchio F, Caprio E, Li J, Evans JD, DeGrandi-Hoffman G, Hamilton M, Chen YP. 2011. Dynamics of persistent and acute deformed wing virus infections in honey bees, Apis mellifera. Viruses 3:2425–2441.
41. Jourdan M, Jousset F-X, Gervais M, Skory S, Bergoin M, Dumas B. 1990. Cloning of the genome of a densovirus and rescue of infectious virions from recombinant plasmid in the insect host Spodoptera littoralis. Virology 179:403–409.

42. Tijssen P, Li Y, El-Far M, Szelei J, Letarte M, Zádori Z. 2003. Organization and expression strategy of the ambisense genome of densonucleosis virus of Galleria mellonella. J Virol 77:10357–10365.

43. Ledermann JP, Suchman IV EL, Black WC, Carlson JO. 2004. Infection and pathogenicity of the mosquito densoviruses AeDNV, HeDNV, and APeDNV in Aedes aegypti mosquitoes (Diptera: Culicidae). J Econ Entomol 97:1828–1835.

44. Ettensohn CA, Wray GA, Wessel GM. 2004. Development of sea urchins, ascidians, and other invertebrate deuterostomes: experimental approaches. Gulf Professional Publishing.

45. Paterson A, Robinson E, Suchman E, Afanasiev B, Carlson J. 2005. Mosquito densonucleosis viruses cause dramatically different infection phenotypes in the C6/36 Aedes albopictus cell line. Virology 337:253–261.
Table 1: Primer/probe sequences, thermocycler parameters, and purpose for each PCR assay

| Purpose                        | Sequence                                                                 |
|--------------------------------|--------------------------------------------------------------------------|
| **AfaDV**                      |                                                                          |
| **Quantitative (qPCR)**        |                                                                          |
| Standard                       | TTCGTAATAGCACCTCGTCACCAGCTAAATATAGATTTTCTCCACACTCTGTATGAAACAGGCAGCTCCACCAGAAGGGATGAGGATAGGATAATTTGCTTA |
| Internal Probe                 | [FAM] TGGATGAAACCAAGCCACGTCACCAGAGTAATTCGGAATGCTGAAAGGAGGATAGGATAATTTGCTTA |
| L-primer position 5664         | CGTAAATAGCACCTCGTCACC                                                   |
| R-primer position 5738         | GACAAATTTACTGGCAGATCATCA                                               |
| **Thermocycling parameters (qPCR):** | 1 cycle of 50°C for 2 min followed by 95°C for 2 min, 50 cycles of 95°C for 15s and 58°C for 1 min – SsoAdvanced Universal Probes Supermix used |
| **AfaDV**                      |                                                                          |
| **Genome Verification (PCR)**  |                                                                          |
| L-primer – 1 position 243      | TTTGAGGTCATATGGGCGGA                                                     |
| R-primer – 1 position 833      | CTTGTCACAACCTCTTCTGCC                                                   |
| L-primer – 2 position 554      | ACCCCACTGATATGCGTA                                                      |
| R-primer – 2 position 1061     | TCCCAAGCTTGGCCAGGA                                                     |
| L-primer – 3 position 831      | TGGGAAGGGAGTTGGA                                                       |
| R-primer – 3 position 1361     | TGCAAAACGCTATCTCTTCTCC                                                 |
| L-primer – 4 position 1266     | GAGCCGATCTGCCAGAGAT                                                    |
| R-primer – 4 position 1702     | TCTCTGACATACCTGGGAGCA                                                  |
| L-primer – 5 Position 1520     | AACAGCAAGACATGGGAGC                                                   |
| R-primer – 5 position 2075     | GATCGGTTGTCGTACATCG                                                  |
| L-primer – 6 position 1979     | GGAGAGCGGACTTGTGATGGAT                                                |
| R-primer – 6 position 2564     | AGAAATTTCTTACCAGCTGAAAGG                                              |
| L-primer – 7 position 2378     | GTGCAGGCTACGTTAAGTTTGT                                                |
| R-primer – 7 Position 2919     | ACAGCAACGGAGTTAGGTTCC                                              |
| L-primer – 8 position 2655     | CCAATTCAAGAGCCTGAGGG                                                 |
| R-primer – 8 position 3144     | AATGTTGCTCCACCAGTTGCA                                                |
| L-primer – 9 position 3066     | CATTGCCAGACATGACAGAGA                                               |
| R-primer – 9 position 3599     | AGTCTGTGGAAACGCTCAG                                                   |
| L-primer – 10 position 3440    | AGCAGAGTCACCACGAGAT                                                   |
| R-primer – 10 position 3895    | CGGTACTGGGATATCCTCTCTGCT                                              |
| L-primer – 11 position 3707    | TGAATCCAAAGTATCGCTCCAG                                              |
| R-primer – 11 position 3999    | ATGAGAGGAGGAGTCGATAGG                                               |
| Primer          | Sequence                                   |
|----------------|--------------------------------------------|
| L-primer – 12  | AGCAGAAGATGATCAGTACCG                     |
| R-primer – 12  | ATTCGCAAAGTGATGGAGGC                      |
| L-primer – 13  | TGGGATTTAGGGAGGAGGT                      |
| R-primer – 13  | AGATCACGTCTAGTAGTGCT                      |
| L-primer – 14  | CACCTTCAGCTGGCGTATA                      |
| R-primer – 14  | TCTTCTCAAGTGATGTCGCA                      |
| L-primer – 15  | TGTGGGCCTTTGAGTGG                        |
| R-primer – 15  | TGTTCGCTGTGTACTTCGT                      |
| L-primer – 16  | TCATCAACATAACAGGCC                      |
| R-primer – 16  | TTTGAGGTCATATGGGCCA                      |

Thermocycling parameters (PCR): 1 cycle 94°C for 2 min, 30 cycles of 45 s at 94°C, 30 s at 56°C and 45 s at 72°C, and final extension for 2 min at 72°C – Taq DNA polymerase used.

| Primer          | Sequence                                   |
|----------------|--------------------------------------------|
| L-primer – 9   | CTTGGGCAGTCAACGAGA                      |
| R-primer – 9   | AGTCTGTGGAACGCTCAG                      |

Thermocycling parameters (PCR): 1 cycle 98°C for 30 s, 35 cycles of 10 s at 98°C, 20 s at 66°C and 20 s at 72°C, and final extension for 2 min at 72°C – Q5 polymerase used.

| Primer          | Sequence                                   |
|----------------|--------------------------------------------|
| L-primer – 252 | CGCaagattATAGAAAAGGCTG                    |
| R-primer – 3187| CGCaagcttCCTAATCCTG                      |

Thermocycling parameters (PCR): 1 cycle 98°C for 30 s, 30 cycles of 10 s at 98°C, 30 s at 66°C and 1 minute 30 s at 72°C, and final extension for 2 min at 72°C – Q5 polymerase used.

| Primer          | Sequence                                   |
|----------------|--------------------------------------------|
| L-primer – 2857| GGGaagcttAGAAACCTAATCC                    |
| R-primer – 5731| CTTATATGTCGGGT                        |

Thermocycling parameters (PCR): 1 cycle 98°C for 30 s, 30 cycles of 10 s at 98°C, 30 s at 67°C and 1 minute 30 s at 72°C, and final extension for 2 min at 72°C – Q5 polymerase used.
| **AfaDV** | L-primer – 1 position 243 | **SSaDV** | SSaDV_NS3_1_F CAATACGCCGATTAGCTTACAG |
| **NS1, NS2, NS3 Cloning** | R-primer – 8 position 3144 | | SSaDV_NS2_1_R_2 TCCCTGCTCACACTAATGTG |
| **(PCR)** | AATGGTGCTCCACCAGTTGC | Thermocycling parameters (PCR): 1 cycle 98°C for 30 s, 30 cycles of 10 s at 98°C, 30 s at 67°C and 1 minute 30 s at 72°C, and final extension for 2 min at 72°C – Q5 polymerase used | Thermocycling parameters (PCR): 1 cycle 98°C for 30 s, 35 cycles of 10 s at 98°C, 30 s at 64°C and 40 s at 72°C, and final extension for 2 min at 72°C – Q5 polymerase used |
Figure 1: Genome architecture and base coverage of *Asterias forbesi* associated densovirus (AfaDV)
(top) Structural hairpins 88nt long located in the inverted terminal repeats at the end of the genome. (middle) Genome organization with ORF colored by putative function. Red corresponds to structural protein (VP), and blue corresponds to non-structural proteins (NS1, NS2, NS3). (bottom) Read coverage distribution across genome. Black line indicates 532x average base coverage across genome.
**Figure 2:** Maximum likelihood phylogeny of densoviruses (AIC; LG +G+I+F). Phylogenetic tree is based on an amino acid alignment performed by MUSCLE of the NS1 region spanning Motif I of the RC endonuclease domain to Motif C of the SF3 helicase domain (amino acid sequence length 437.7 ± 50; mean ± SD). Branch support bootstrapped at 100 iterations and are shown through colored branches. Black branches indicate <80% support, blue branches indicate 80-90% support, and green branches indicate 90-100% support. Terminal node colors correspond to densovirus genus. Italicized names correspond to animal genus and species for which the densovirus was isolated. AfaDV denoted in bold with *.
Figure 3: AfaDV prevalence among sea star populations along the North American Atlantic Coast. A total of 134 animals from three species of sea stars were screened via qPCR or PCR for AfaDV. Year of sampling shown in parentheses. Prevalence for each species corresponds to the number of animals positive for AfaDV divided by the total listed next to each species.
Figure 4: Viral load and tissue prevalence of AfaDV. (left) Pearson's correlation between viral load and animal length reported as total diameter. Colors correspond to sample type. Black dots represent cross-section samples; (middle) Viral load comparison between tissue type. (right); Prevalence of AfaDV among tissue types. *** p = < 0.001. NS = No significance.