Vol. 105: 237–242, 2013
doi: 10.3354/dao02628

DISEASES OF AQUATIC ORGANISMS
Dis Aquat Org

Published September 3

FREE ACCESS

Metagenomic identification of a nodavirus and a circular ssDNA virus in semi-purified viral nucleic acids from the hepatopancreas of healthy *Farfantepenaeus duorarum* shrimp

Terry Fei Fan Ng[1,6,*], Shankar Alavandi[2], Arvind Varsani[3,4,5], Scott Burghart[1], Mya Breitbart[1,*]

[1]University of South Florida, College of Marine Science, St. Petersburg, Florida, USA
[2]Central Institute of Brackishwater Aquaculture, #75, Santhome High Road, RA Puram, Chennai, India
[3]School of Biological Sciences, University of Canterbury, Christchurch 8140, New Zealand
[4]Biomolecular Interaction Centre, University of Canterbury, Christchurch 8140, New Zealand
[5]Electron Microscope Unit, Division of Medical Biochemistry, Department of Clinical Laboratory Sciences, University of Cape Town, Observatory 7700, South Africa
[6]*Present address:* Blood Systems Research Institute, San Francisco, California, USA

ABSTRACT: Fisheries and aquaculture are impacted sporadically by newly emerged viral diseases. In most cases, searches for a causative pathogen only occur after a serious disease has emerged. As random shotgun sequencing (metagenomics) offers opportunities to identify novel viruses preemptively, the method was tested on nucleic acids extracted from the hepatopancreas of 12 healthy northern pink shrimp *Farfantepenaeus duorarum* captured from the Gulf of Mexico. Among the sequences, a nodavirus (*Farfantepenaeus duorarum* nodavirus, *Fd*NV) and a virus with similarities to circoviruses and cycloviruses that possess circular single-stranded DNA (ssDNA) genomes, were identified. The *Fd*NV genome sequence was most closely related phylogenetically to nodaviruses causing white tail disease in *Macrobrachium rosenbergii* and muscle necrosis disease in *Litopenaeus vannamei*. While the circular ssDNA virus represents the third to be detected in association with a marine invertebrate, transmission trials are needed to confirm its infectivity for *F. duorarum*. This study highlights the potential for using metagenomic approaches in fisheries and aquaculture industries to identify new potential pathogens in asymptomatic marine invertebrates, uncharacterized pathogens causing a new disease, or multiple pathogens associated with disease syndromes.

KEY WORDS: Fisheries · Pathogen discovery · Metagenomics · Nodavirus · Circovirus

INTRODUCTION

In 2010, global aquaculture of over 600 species including finfish, crustaceans, and mollusks produced approximately 60 million tons of seafood with an estimated value of US $119 billion (FAO 2011). However, aquaculture worldwide has become increasingly susceptible to new diseases due mainly to movement of live animals, the increased use of high-density farming systems, and increasing anthropogenic pressures on aquatic ecosystems (FAO 2011). With tropical shrimp aquaculture, it has been esti-

*Corresponding authors.
Emails: terryfeifan@gmail.com, mya@marine.usf.edu

mated that up to 40% of annual production (>$3 billion) is currently being lost, mainly due to viral diseases (Stentiford et al. 2012).

Although emerging diseases often cause substantial economic losses, little effort is generally expended on preemptively identifying potential pathogens before they cause a problem to aquaculture. This delays development of molecular diagnostic tests for identifying pathogen reservoirs and vectors, host and geographic ranges, modes of transmission, and genetic relatedness to other pathogens, impeding our understanding of disease epidemiology and ability to minimize disease impacts (Walker & Winton 2010). The reactive approach also neglects the fact that many potential pathogens have evolved to coexist in symbiosis with particular hosts, with transition to disease states only occurring after some ecological imbalance promoted by environmental or anthropogenic factors (Walker & Winton 2010). Understanding which potential pathogens exist naturally in commercially important species would thus assist preparedness for preventing or managing disease should a pathogen emerge as a problem.

Various methods are available for identifying and characterizing novel microorganisms. Metagenomic sequencing of nucleic acids extracted from fluids or tissue homogenates filtered or enriched to concentrate and semi-purify virus particles has proved useful in identifying genome sequences of novel viruses in humans and various domestic and wild animals (Breitbart et al. 2002, Allander et al. 2005, Delwart 2007, Victoria et al. 2008, Ng et al. 2009, 2011a,b, 2012, Svraka et al. 2010). Once sequence data are available, options abound to develop sensitive, specific, and rapid molecular tests for virus diagnosis and epidemiology.

Due to the wild shrimp fishery in southern parts of the Gulf of Mexico showing signs of being depleted, aquaculture of northern pink shrimp *Farfantepenaeus duorarum* has become increasingly important to the economy of the region (Arreguin-Sanchez et al. 2008). To avoid disease impacts, pathogen-free breeding stocks of *F. duorarum* have also been sought to support the industry (Samocha et al. 2008). To investigate whether viruses might exist in healthy *F. duorarum* indigenous to the Gulf of Mexico that could pose a potential risk to aquaculture, random shotgun sequencing was undertaken on extracts of hepatopancreas tissue. This led to the identification of a new nodavirus and a virus with similarities to circoviruses and cycloviruses that possess circular single-stranded DNA (ssDNA) genomes.

MATERIALS AND METHODS

Overtly healthy juvenile *Farfantepenaeus duorarum* (6–8 g) caught near Tarpon Springs, Florida, USA, in the Gulf of Mexico, were transported to the laboratory in aerated buckets. Hepatopancreas tissue dissected aseptically from 12 shrimp was pooled and homogenized in sterile SM buffer (50 mM Tris, 10 mM $MgSO_4$, 0.1 M NaCl, pH 7.5) using a Tissumizer (Tekmar). Virus particles were purified from filtered homogenate as described previously (Breitbart & Rohwer 2005, Ng et al. 2011a, 2012). Briefly, homogenate was clarified by centrifugation at 10 000 × *g* (10 min), filtered through a 0.22 µm filter, mixed with 0.2 volume chloroform for 10 min, and then incubated with 2.5 U DNase I and 0.25 U RNase A per µl at 37°C for 3 h. DNA and RNA were extracted from purified material using the Qiagen DNeasy Blood & Tissue Kit and RNeasy Mini Kit, respectively.

RNA and DNA were sequenced using cDNA synthesis, DNA amplification, cloning, and sequencing methods described previously (Breitbart & Rohwer 2005, Ng et al. 2011a). Briefly, DNA was amplified using Phi29 DNA polymerase (Genomiphi, GE Healthcare), fragmented and amplified again using a Whole Genome Amplification kit (Sigma-Aldrich). RNA was converted to cDNA and amplified by PCR using a TransPlex Whole Transcriptome Amplification kit (Sigma-Aldrich). Randomly amplified DNA libraries were cloned into a TOPO TA vector (Invitrogen) and transformed into competent cells. Inserts in 50 to 130 clones from the DNA and RNA libraries were amplified by PCR using M13 primers, and amplicons >150 bp were Sanger sequenced, trimmed, and assembled into contigs applying a match size = 35, minimum match percentage = 95% using the SeqMan Pro-assembler (DNASTAR). Singlets and assembled contigs were compared to sequences in the GenBank non-redundant database using BLASTX (Altschul et al. 1990, 1997), leading to the identification of a circovirus-like DNA sequence and a nodavirus-like RNA sequence.

To generate a complete circovirus-like genome sequence by inverse PCR (Ng et al. 2009, 2011a), DNA was amplified randomly using Phi29 DNA polymerase (TempliPhi, GE Healthcare) before PCR for 45 cycles employing the outward facing primers 5′^{124}-TGA CAT TGG GAT ACC ACT GG^{143}-3′ and 5′^{126}-TCA AGG ATA CTG CTG CCA TG^{107}-3′. The ~2 kb DNA amplified by PCR was cloned and Sanger sequenced by primer walking.

Amino acid sequences of the circovirus-like replication initiator protein (Rep) and nodavirus capsid protein (Cap) were aligned to various homologues

using MUSCLE (Edgar 2004) with manual editing. The multiple alignments were used to infer maximum-likelihood phylogenetic trees using PHYML Version 3.0 (Guindon et al. 2010) with approximate likelihood-ratio test (aLRT) branch support (Anisimova & Gascuel 2006) and an LG model of substitution. Mesquite (version 2.75) was used to collapse branches with <60% aLRT branch support.

RESULTS AND DISCUSSION

Virus particles were partially purified from a homogenate of hepatopancreas tissue from 12 wild *Farfantepenaeus duorarum* using filtration, chloroform, and nuclease treatment. Sequence analysis of clones derived from either RNA or DNA extracted and amplified from this material identified a new DNA virus possessing a circular genome designated shrimp hepatopancreas-associated circular DNA virus (ShrimpCDV) and a new nodavirus designated *Farfantepenaeus duorarum* nodavirus (*Fd*NV).

ShrimpCDV

The 1956 nt ssDNA genome of ShrimpCDV (Fig. 1A; GenBank Accession KC441518) contains 2 genes encoding putative Rep and Cap proteins transcribed bi-directionally. A putative DNA hairpin structure containing a sequence (AGG TAT TAC) similar to the conserved nonanucleotide motif of circoviruses exists in the short intergenic region (SIR; Fig. 1A). In pairwise distance analyses, the ShrimpCDV Rep protein was 21 to 34% identical to the cognate protein of circoviruses, cycloviruses, and other unclassified ssDNA viruses. Fig. 1B illustrates the phylogenetic relationships between ShrimpCDV and other circular ssDNA viruses. According to the classification scheme based on genome organization developed recently for circular ssDNA viruses (Rosario et al. 2012), ShrimpCDV is a Type IV virus.

Numerous novel circular ssDNA virus genomes have been identified recently by viral metagenomics from insects, fecal matter, and various environmental samples, including, for example, reclaimed water and sewage (Delwart & Li 2012, Rosario et al. 2012). Although many genomes have been identified in seawater and one was found associated with fish (Rosario et al. 2009, Lorincz et al. 2011), with regard to marine invertebrates, circular ssDNA viruses have only recently been found associated with 2 copepod species (*Acartia tonsa* and *Labidocera aestiva*; Dunlap et al. 2013). Due to their prevalence in the environment, it is not unexpected that circular ssDNA viruses should be found associated with marine organisms. Since the hepatopancreas is a digestive organ of shrimp, and semi-purified viral DNA extracted from the hepatopancreas tissue was amplified using Phi29 DNA polymerase followed by random PCR, it cannot be discounted that ShrimpCDV represents an ingested environmental contaminant resilient to degradation by the shrimp digestive system. To gain more knowledge regarding the host of ShrimpCDV, molecular screening of wild *Farfantepenaeus duorarum* needs to be undertaken to verify the presence of this virus in shrimp tissues exclusive of the digestive track, and bioassays should be performed to test its ability to infect *F. duorarum*.

*Fd*NV

Two nodavirus-like sequences were identified among clones generated from the *Farfantepenaeus duorarum* hepatopancreas RNA virome. Nodaviruses are small, spherical, non-enveloped viruses containing a genome comprised of 2 positive sense ssRNAs (Scherer & Hurlbut 1967, Scotti et al. 1983, Dasmahapatra et al. 1985, Reinganum et al. 1985, Zeddam et al. 1999), with the larger RNA1 encoding a RNA-dependent RNA polymerase and the smaller RNA2 encoding the structural capsid protein (Friesen & Rueckert 1981, Ball 1995, Nagai & Nishizawa 1999). The *Nodaviridae* includes insect-infecting viruses classified in the genus *Alphanodavirus* (Scherer & Hurlbut 1967, Scotti et al. 1983, Dasmahapatra et al. 1985, Reinganum et al. 1985, Zeddam et al. 1999), fish-infecting viruses classified in the genus *Betanodavirus* (Munday et al. 2002), and the as yet unclassified *Macrobrachium rosenbergii* nodavirus (*Mr*NV) and *Penaeus vannamei* nodavirus (*Pv*NV; from *Litopenaeus vannamei*) that cause muscle diseases in these shrimp species (Arcier et al. 1999, Tang et al. 2007).

In our study, 2 clones were identified from the *Farfantepenaeus duorarum* hepatopancreas tissue that shared amino acid identities to previously described nodaviruses. One clone contained a 403 nt insert encoding an RNA-dependent RNA polymerase (RdRp) partial sequence (GenBank Accession: KC441519) and the other contained a 236 nt insert encoding a capsid protein partial sequence (GenBank Accession: KC441520). Phylogenetic analysis using the partial sequence of the *Fd*NV capsid protein, which is used to classify nodaviruses, revealed 43 to 51% amino acid identity to the 2 previously characterized

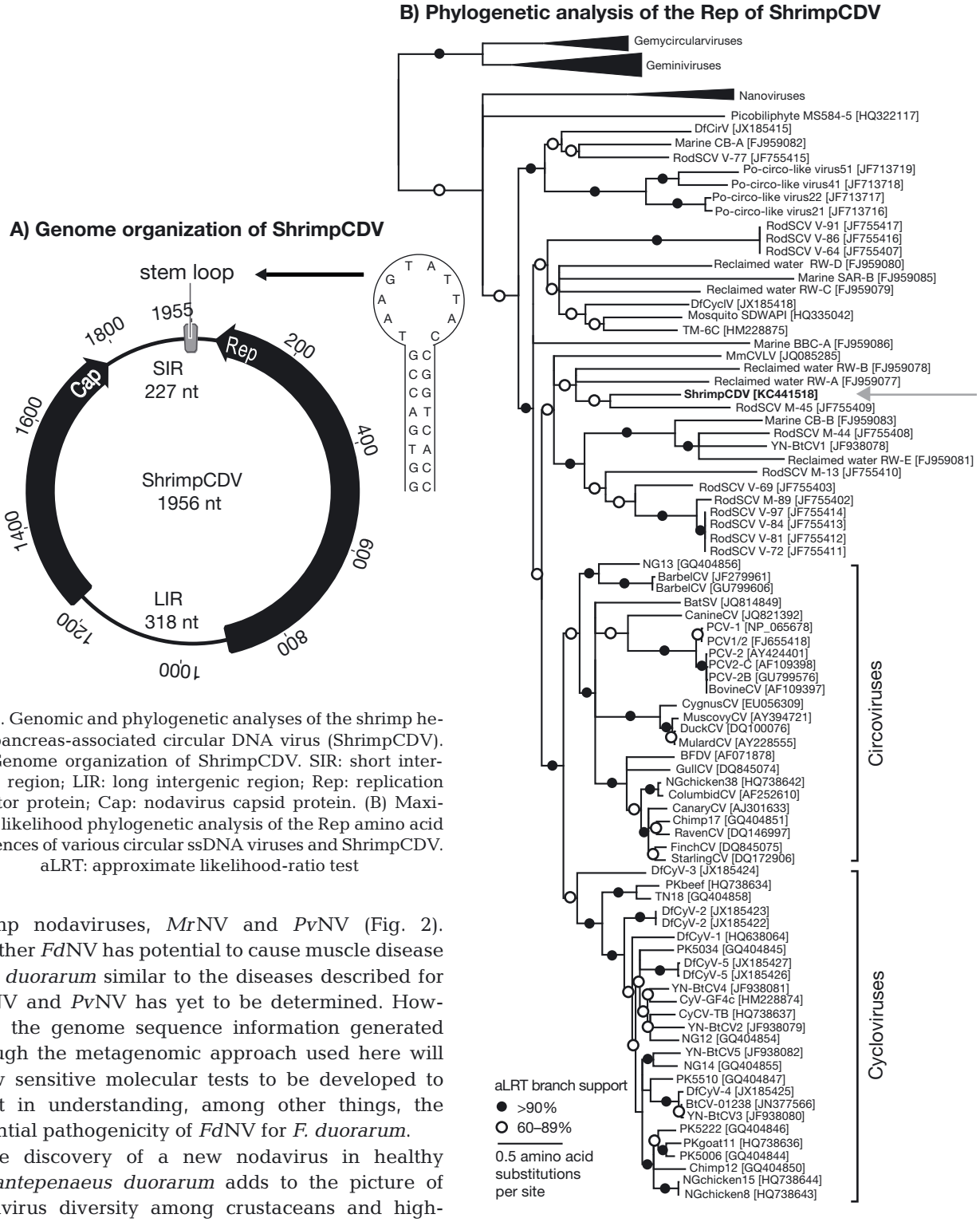

Fig. 1. Genomic and phylogenetic analyses of the shrimp hepatopancreas-associated circular DNA virus (ShrimpCDV). (A) Genome organization of ShrimpCDV. SIR: short intergenic region; LIR: long intergenic region; Rep: replication initiator protein; Cap: nodavirus capsid protein. (B) Maximum likelihood phylogenetic analysis of the Rep amino acid sequences of various circular ssDNA viruses and ShrimpCDV. aLRT: approximate likelihood-ratio test

shrimp nodaviruses, *Mr*NV and *Pv*NV (Fig. 2). Whether *Fd*NV has potential to cause muscle disease in *F. duorarum* similar to the diseases described for *Mr*NV and *Pv*NV has yet to be determined. However, the genome sequence information generated through the metagenomic approach used here will allow sensitive molecular tests to be developed to assist in understanding, among other things, the potential pathogenicity of *Fd*NV for *F. duorarum*.

The discovery of a new nodavirus in healthy *Farfantepenaeus duorarum* adds to the picture of nodavirus diversity among crustaceans and highlights the risks of translocation of unknown viruses of live crustaceans. For example, *Mr*NV and *Pv*NV now occur in both Western and Eastern Hemisphere regions where these species are cultured as the result of trade in live crustaceans (Qian et al. 2003, Yoganandhan et al. 2006). The low genetic relatedness of *Fd*NV to *Mr*NV and *Pv*NV (Fig. 2) suggests that these viruses have evolved over long periods in their preferred host species, and that metagenomic approaches such as those used here might find similarly divergent nodaviruses in other shrimp species.

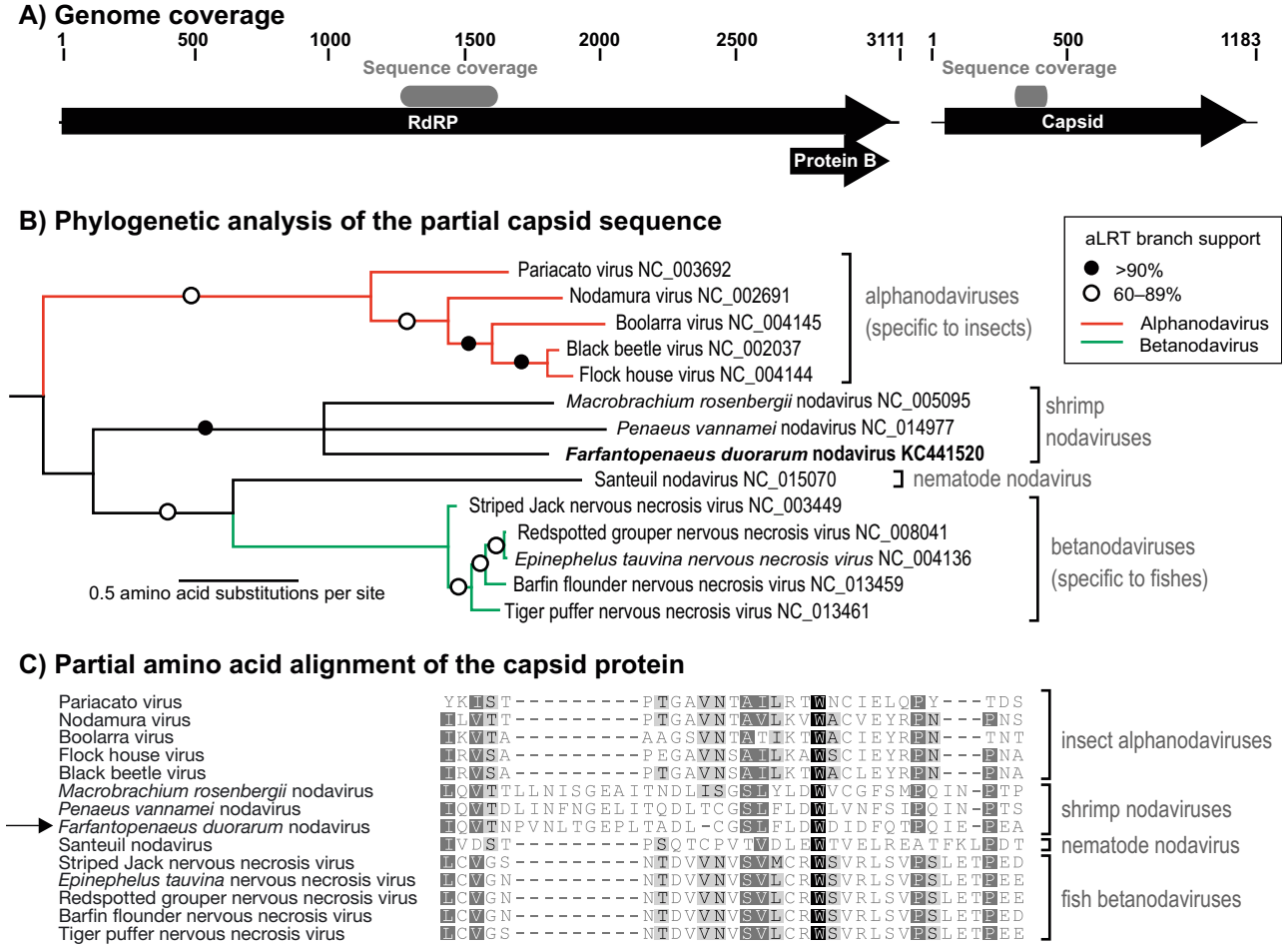

Fig. 2. Genomic and phylogenetic analyses of the *Farfantepenaeus duorarum* nodavirus (*Fv*NV). (A) Sequence coverage of *Fv*NV, indicated by grey boxes. (B) Maximum likelihood phylogenetic analysis of the partial amino acid sequences of the capsid protein from *Fv*NV and representatives of the *Nodaviridae*. (C) Partial amino acid alignment of the capsid protein of the *Fv*NV with other nodaviruses

Metagenomics

The use of a metagenomic approach to identify a new nodavirus and a novel circular ssDNA virus in overtly healthy *Farfantepenaeus duorarum* highlights the potential value of such methods for discovering unknown viruses in commercially-important fisheries and aquaculture species. Such preemptive discoveries offer opportunities to develop sensitive molecular diagnostic tests and to guide research to determine the potential threat of newly-discovered viruses in emerging as a significant disease (Walker & Winton 2010). Metagenomic approaches can similarly be used to identify undiagnosed pathogens causative of newly-emerged diseases (Allander et al. 2005, Delwart 2007, Victoria et al. 2008, Ng et al. 2009, 2011a, Svraka et al. 2010), thereby expediting development of diagnostic tools to help prevent or manage these diseases.

Acknowledgements. This work was funded through grants DEB-1239976 and OCE-1049670 from the National Science Foundation to M.B. T.N. was funded by the William and Elsie Knight Oceanographic Fellowship. S.V.A. was supported by the DBT Overseas Fellowship from the Department of Biotechnology, Government of India. We thank B. Dwivedi for assistance with bioinformatics, S. Habtes and E. Peebles for identification of shrimp, and T. MacDonald for scientific discussion.

LITERATURE CITED

Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B (2005) Cloning of a human parvovirus by molecular screening of respiratory tract samples. Proc Natl Acad Sci USA 102:12891–12896

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

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search

programs. Nucleic Acids Res 25:3389–3402
Anisimova M, Gascuel O (2006) Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst Biol 55:539–552
Arcier JM, Herman F, Lightner DV, Redman RM, Mari J, Bonami JR (1999) A viral disease associated with mortalities in hatchery-reared postlarvae of the giant freshwater prawn *Macrobrachium rosenbergii.* Dis Aquat Org 38:177–181
Arreguin-Sanchez F, Zetina-Rejón M, Ramírez-Rodríguez M (2008) Exploring ecosystem-based harvesting strategies to recover the collapsed pink shrimp (*Farfantepenaeus duorarum*) fishery in the southern Gulf of Mexico. Ecol Model 214:83–94
Ball LA (1995) Requirements for the self-directed replication of flock house virus RNA 1. J Virol 69:720–727
Breitbart M, Rohwer F (2005) Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. Biotechniques 39:729–736
Breitbart M, Salamon P, Andresen B, Mahaffy JM and others (2002) Genomic analysis of uncultured marine viral communities. Proc Natl Acad Sci USA 99:14250–14255
Dasmahapatra B, Dasgupta R, Ghosh A, Kaesberg P (1985) Structure of the black beetle virus genome and its functional implications. J Mol Biol 182:183–189
Delwart EL (2007) Viral metagenomics. Rev Med Virol 17: 115–131
Delwart E, Li L (2012) Rapidly expanding genetic diversity and host range of the *Circoviridae* viral family and other Rep encoding small circular ssDNA genomes. Virus Res 164:114–121
Dunlap DS, Ng TF, Rosario K, Barbosa GJ, Greco AM, Breitbart M, Hewson I (2013) Molecular and microscopic evidence of viruses in marine copepods. Proc Natl Acad Sci USA 110:1375–1380
Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797
FAO (Food and Agriculture Organization of the United Nations) (2011) The state of the world fisheries and aquaculture 2010. FAO, Rome
Friesen PD, Rueckert RR (1981) Synthesis of black beetle Virus proteins in cultured *Drosophila* cells: differential expression of RNAs 1 and 2. J Virol 37:876–886
Guindon S, Dufayard JF, 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
Lorincz M, Cságola A, Farkas SL, Székely C, Tuboly T (2011) First detection and analysis of a fish circovirus. J Gen Virol 92:1817–1821
Munday BL, Kwang J, Moody N (2002) Betanodavirus infections of teleost fish: a review. J Fish Dis 25:127–142
Nagai T, Nishizawa T (1999) Sequence of the non-structural protein gene encoded by RNA1 of striped jack nervous necrosis virus. J Gen Virol 80:3019–3022
Ng TFF, Suedmeyer WK, Gulland F, Wheeler E, Breitbart M (2009) Novel anellovirus discovered from a mortality event of captive California sea lions. J Gen Virol 90: 1256–1261
Ng TFF, Wheeler E, Greig D, Waltzek TB, Gulland F, Breitbart M (2011a) Metagenomic identification of a novel anellovirus in Pacific harbor seal (*Phoca vitulina richardsii*) lung samples and its detection in samples from multiple years. J Gen Virol 92:1318–1323
Ng TFF, Willner DL, Lim YW, Schmieder R and others (2011b) Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. PLoS ONE 6: e20579
Ng TF, Marine R, Wang C, Simmonds P and others (2012) High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. J Virol 86: 12161–12175
Qian D, Shi Z, Zhang S, Cao Z and others (2003) Extra small virus-like particles (XSV) and nodavirus associated with whitish muscle disease in the giant freshwater prawn, *Macrobrachium rosenbergii*. J Fish Dis 26:521–527
Reinganum C, Bashiruddin JB, Cross GF (1985) Boolarra virus: a member of the *Nodaviridae* isolated from *Oncopera intricoides* (Lepidoptera: Hepialidae). Intervirology 24:10–17
Rosario K, Duffy S, Breitbart M (2009) Diverse circovirus-like genome architectures revealed by environmental metagenomics. J Gen Virol 90:2418–2424
Rosario K, Duffy S, Breitbart M (2012) A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. Arch Virol 157:1851–1871
Samocha TM, Gandy RL, Morris TC, Patnaik S and others (2008) Development of viral pathogen free broodstock populations of the Atlantic pink shrimp *Farfantepenaeus duorarum* and the Atlantic white shrimp *Litopenaeus setiferus*. In: Proc Aquaculture America 2008. World Aquaculture Society Meeting, Lake Buena Vista, FL, p 399 (Abstract)
Scherer WF, Hurlbut HS (1967) Nodamura virus from Japan: a new and unusual arbovirus resistant to diethyl ether and chloroform. Am J Epidemiol 86:271–285
Scotti PD, Dearing S, Mossop DW (1983) Flock House virus: a nodavirus isolated from *Costelytra zealandica* (White) (Coleoptera: Scarabaeidae). Arch Virol 75:181–189
Stentiford GD, Neil DM, Peeler EJ, Shields JD and others (2012) Disease will limit future food supply from the global crustacean fishery and aquaculture sectors. J Invertebr Pathol 110:141–157
Svraka S, Rosario K, Duizer E, van der Avoort H, Breitbart M, Koopmans M (2010) Metagenomic sequencing for virus identification in a public-health setting. J Gen Virol 91:2846–2856
Tang KFJ, Pantoja CR, Redman RM, Lightner DV (2007) Development of *in situ* hybridization and RT-PCR assay for the detection of a nodavirus (*Pv*NV) that causes muscle necrosis in *Penaeus vannamei*. Dis Aquat Org 75: 183–190
Victoria JG, Kapoor A, Dupuis K, Schnurr DP, Delwart EL (2008) Rapid identification of known and new RNA viruses from animal tissues. PLoS Pathog 4:e1000163
Walker PJ, Winton JR (2010) Emerging viral diseases of fish and shrimp. Vet Res 41:51
Yoganandhan K, Leartvibhas M, Sriwongpuk S, Limsuwan C (2006) White tail disease of the giant freshwater prawn *Macrobrachium rosenbergii* in Thailand. Dis Aquat Org 69:255–258
Zeddam JL, Rodriguez JL, Ravallec M, Lagnaoui A (1999) A noda-like virus isolated from the sweetpotato pest *Spodoptera eridania* (Cramer) (Lep.; Noctuidae). J Invertebr Pathol 74:267–274

Editorial responsibility: Craig Carlson, Santa Barbara, California, USA

Submitted: October 12, 2012; Accepted: June 11, 2013
Proofs received from author(s): July 31, 2013