

Detection and molecular characterization of a novel piscine-myocarditis-like virus from baitfish in the USA

Sunil K. Mor^{1,2} · Nicholas B. D. Phelps^{1,2}

Received: 22 November 2015 / Accepted: 23 April 2016 / Published online: 3 May 2016
© Springer-Verlag Wien 2016

Abstract During a survey of apparently healthy baitfish from the state of Minnesota, a novel piscine-myocarditis-like virus (PMCLV) was detected in golden shiners (*Notemigonus crysoleucas*). The nearly complete genome sequence is 5819 nt long, including a partial 5' untranslated region (UTR) of 100 nt. The sequence is divided into three ORFs: the complete ORF1 and ORF2, encoding proteins of 818 and 831 amino acids, respectively, and a partial ORF3 encoding 248 amino acids of the corresponding protein. This novel virus sequence was most closely related to piscine myocarditis virus (PMCV), with a 49.0 % and 58.2 % amino acid identity in the ORF1 (YP005481249)- and ORF2 (YP004581250)-encoded proteins, respectively. Six of 56 retail outlets (e.g., bait shops) were positive during the 2014–2015 survey, indicating a 10.7 % prevalence of the novel virus in this population of golden shiners. Currently, there is no disease that is known to be associated with this virus.

Introduction

Baitfish farming is a large aquaculture industry in the United States, with 257 baitfish farms producing farm-gate value of \$38 million in 2005 and retail sales 10–15 times greater [17]. The baitfish industry in the USA distributes more than 10 billion fish per year across the country. These

fish are distributed through wholesale and retail networks to recreational anglers who take the live baitfish to rivers and lakes, where they may be consumed by predators or released into the wild [6]. In addition, large quantities of baitfish are sold as forage for the production of predatory species such as walleye and muskellunge. The golden shiner (*Notemigonus crysoleucas*) is one of the most important baitfish in the USA, with 76 farms producing \$17.1 million worth of these fish in 2005 [17]. The majority of these golden shiners routinely undergo regulatory inspection for important viral diseases, such as viral hemorrhagic septicemia virus. However, the health status of fish at retail remains largely unknown. To identify potential pathogens that are being transported via the baitfish trade, a survey of golden shiners at retail outlets (e.g., bait shops) in Minnesota was performed. A metagenomics approach was used for the detection of all possible viruses associated with these apparently healthy golden shiners. We detected a previously uncharacterized novel piscine-myocarditis-like virus in golden shiners, which is the subject of this study.

Piscine myocarditis virus (PMCV) is a non-enveloped icosahedral double-stranded RNA virus belonging to the family *Totiviridae*. Its 50-nm virion is slightly larger than that of other members of family *Totiviridae* [7], which has five recognized genera, *Totivirus*, *Victorivirus*, *Leishmaniavirus*, *Giardiavirus*, and *Trichomonasvirus*. The members of *Totivirus* and *Victorivirus* infect yeast and fungi, while members of *Giardiavirus* and *Leishmaniavirus* infect parasitic protozoa [18]. Two unassigned members of the family *Totiviridae* are infectious myonecrosis virus (IMNV) and PMCV. These two unassigned viruses cluster with members of the genus *Giardiavirus*, but it is not clear if they should be regarded as forming a new genus within the family *Totiviridae* or a new species within the genus

✉ Sunil K. Mor
kumars@umn.edu

¹ Minnesota Veterinary Diagnostic Laboratory, 1333 Gortner Avenue, St. Paul, MN 55108, USA

² Department of Veterinary Population Medicine, University of Minnesota, 1365 Gortner Avenue, St. Paul, MN 55108, USA

Giardiavirus. Other unclassified viruses in the family *Totiviridae* include *Armigeres subalbatus* totivirus (AsTV) of mosquitoes (*Armigeres subalbatus*), *Drosophila* totivirus (DTV) of fruit flies (*Drosophila melanogaster*), Omono River virus (OMRV) of *Culex* mosquitoes and Tianjin totivirus (ToV-TJ) of bats (*Myotis ricketti*) [8, 10, 20, 21]. A new genus named *Artivirus* has been proposed to include three arthropod-infecting totiviruses (AsTV, DTV and IMNV) [21].

PMCV is the causative agent of cardiomyopathy syndrome (CMS), which causes inflammatory heart disease, primarily in farmed and wild Atlantic salmon (*Salmo salar* L.) [5, 7, 16, 19]. The disease was first detected in Norway in 1985 and later was diagnosed in farmed Atlantic salmon in the Faeroe Islands in Scotland and in British Columbia, Canada [1, 2, 12, 13]. Recently, PMCV was detected in Atlantic argentine (*Argentina silus*, Ascanius) in Norway [3, 7].

The genome of PMCV is 6688 nucleotides in length and contains three open reading frames (ORF1, ORF2 and ORF3) [7]. Open reading frame 1 encodes a putative capsid protein of 861 amino acids, while ORF2 encodes an RNA-dependent RNA polymerase (RdRp) and is the only region of the genome identifiable by BLASTx. The putative ORF3 (302 aa) has not been described previously in any member of the family *Totiviridae*, and the precise role of the encoded product is unknown [15, 19]. In this study, we describe a novel PMCV-like virus in golden shiner, named GS-PMCLV (golden shiner piscine-myocarditis-like virus) from Minnesota, USA.

Materials and methods

Source of samples

Thirty golden shiners were collected, using a dipnet, from 56 retail outlets in Minnesota from May to October 2014 (n = 25) and December 2014 to February 2015 (n = 31). All collected fish were delivered live on the same day to the University of Minnesota Veterinary Diagnostic Laboratory (UMVDL). Upon arrival, fish were euthanized with an overdose of MS-222 (Argent Chemical Laboratories, WA, USA), and a gross examination and necropsy were performed.

Virus isolation

The entire viscera of 30 fish were pooled in groups of five, resulting in six pooled samples per location. Virus isolation was performed according to the US Fish and Wildlife Service and American Fisheries Society – Fish Health Section Blue Book as described previously [9, 11]. Briefly,

a 10 % homogenate of the entire viscera was prepared in Hanks' balanced salt solution followed by centrifugation at $2500 \times g$ for 15 min at 4 °C. The supernatant was inoculated onto epithelioma papulosum cyprini (EPC), fathead minnow (FHM), and chinook salmon embryo (CHSE) cell lines at 15 and 22 °C. If no cytopathic effects (CPE) were observed after 14 days of incubation, the cell cultures were passaged on fresh cells and observed for an additional 14 days.

Illumina sequencing

The clarified supernatant of processed tissue pools from 25 retail outlets collected during May – October 2014 were combined into one pool per location for next-generation sequencing (NGS) as described previously [9]. Briefly, good-quality total viral RNA free from ribosomal RNA and DNA was extracted using TRIzol LS Reagent (Invitrogen, NY, USA), followed by RNA purification using a QIAamp Viral RNA Mini Kit (QIAGEN, Valencia, CA). Extracted RNA was submitted to the University of Minnesota Genomic Center (UMGC) for Illumina Miseq analysis. The library was prepared using TruSeq RNA v2 followed by sequencing on a 300 PE run on the MiSeq. The obtained sequence reads were analyzed using CLC Genomics Workbench 7.5 (<http://www.clcbio.com>). After trimming adaptor sequences and sequence quality testing, contigs were prepared by *de novo* assembly. Extracted contigs were analyzed using BLASTx at NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The ORFs were predicted using the ORF Finder tool (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>).

Sequence analysis and phylogeny

Based on the study sequence, sets of overlapping primers were designed for confirmation of the Illumina sequence. Amplified PCR products were purified and sequenced in both directions using the same forward and reverse primers. The nucleotide sequences generated in this study were aligned and compared with previously reported sequences of PMCV and other members of the family *Totiviridae* available in the GenBank database. Sequences were aligned using ClustalW in MEGA 6.06 software [14]. The best substitution model for analysis of protein sequences by the maximum-likelihood method was selected based on the lowest BIC (Bayesian information criterion) score in MEGA 6.05. The amino acid substitution model LG+G (gamma distribution with 5 rate categories) was used for ORF1 and ORF2 to generate phylogenetic trees with 1000 bootstrap replicates. Geneious Pro software was used for sequence comparison and calculation of pairwise identity of nucleotide (nt) and amino acid (aa) sequences [4].

Reverse transcription polymerase chain reaction (RT-PCR)

Based on the study sequence, a primer set was designed targeting the RdRp gene for standardization of an RT-PCR assay to screen all locations (one pooled sample per location, $n = 56$) in the survey. The primers UM-GS-PMCLV_RdRp_F (5'-TATTGAGCGGGTGGAGATTC-3') and UM-GS-PMCLV_RdRp_R (5'-TAGGCGGATGTACCCACTTC-3') were designed targeting nt 3701-3720 and 4602-4622 of the RdRp gene, generating an amplified product of 921 bp. A OneStep RT-PCR Kit (QIAGEN, Valencia, CA) was used for the PCR reaction, and a 25- μ L reaction mix was prepared according to the instructions provided with the kit. The amplified PCR products were analyzed in an ethidium-bromide-stained 1.2 % agarose gel. A single band of the expected product size confirmed the presence of target virus. Amplified products were purified using a QIAGEN PCR purification kit and then submitted to the UMGC for Sanger sequencing using the forward and reverse primers that were used for RT-PCR. These forward and reverse sequences were aligned using Sequencher 5.1 software (<http://www.genecodes.com>), followed by BLAST analysis (<http://www.ncbi.nlm.nih.gov>).

Results

Gross examination and virus isolation

No clinical abnormalities or notable lesions were observed in the golden shiners collected from 56 retail outlets as part of the survey. No CPE associated with the novel GS-PMCLV was observed after two passages on EPC, FHM, or CHSE cells at 15 or 22 °C.

Sequence analysis and phylogeny

In three out of 25 samples, *de novo* assembly resulted in sequence contigs belonging to a novel virus resembling members of the family *Totiviridae*. Sequence quality was confirmed by Sanger sequencing, which gave 100 % nt sequence identity to the Illumina sequence from nt 1-5819. The sequences were most closely related to the previously reported PMCV (accession no. HQ339954; protein accession nos. YP004581249 [ORF1], YP004581250 [ORF2], and YP004581251 [ORF3]) isolated from Atlantic salmon. The nearly complete genome sequence of the novel virus is 5819 nt long, which is 869 nt shorter than the reference PMCV genome sequence (6688 nt). The 5' untranslated regions (UTRs) in the novel sequence is 100 nt long,

smaller than the 444-nt 5' UTR of the PMCV reference (Table 1). This indicates that the novel sequence likely has an incomplete 5' UTR. The sequence, named GS-PMCLV/USA/MN/2014, has been submitted to GenBank with accession number KT725636.

The novel sequence is divided into three ORFs (ORF1, ORF2 and ORF3) (Fig. 1). ORF1 encodes a putative structural protein of 818 aa from nt positions 101 to 2557 (2457 nt) with a start codon at nt 101-103 and a stop codon at nt 2555-2557. The predicted product of ORF1 is 43 aa shorter than the 861-aa ORF1 product of the PMCV reference genome sequence. The ORF1-encoded protein has a predicted molecular mass of 88.2 kDa, which is smaller than the 91.8-kDa protein of the reference strain, but the pI value of 6.85 is higher than 5 that of the reference strain (pI = 5.4). The GC content of 51.6 % is also less than that of the reference strain (56.2 %). Similar to reference strain, the sequence from this study does not have a 2A-like motif or a dsRNA-binding motif, which are characteristic features of IMNV.

ORF2 has interesting features that are different from those of the reference PMCV genome. It encodes an RdRp similar to those of PMCV and other members of the family *Totiviridae*. This protein is 831 aa long, which is 105 aa longer than that encoded by ORF2 in the reference genome (726 aa). ORF2 is present in reading frame +1 at nucleotide position 2305-4800 with an overlap of 252 nt with ORF1, which is different from the reference PMCV genome, where ORF2 is present in reading frame +3 with a gap of 87 nt between it and the stop codon of ORF1. The GC content is 40.7 %, compared to 43.56 % for the

Table 1 Comparison of the genome sequence of golden shiner-piscine myocarditis-like virus GS-PMCLV with the reference PMCV sequence

Genome features	GS-PMCLV	PMCV AL V-708
Total genome size	5819 nucleotides*	6688 nucleotides
5' UTR	100 nucleotides*	444 nucleotides
3' UTR	-	241 nucleotides
ORF 1		
Size	818 amino acids	861 amino acids
Mass	88.2 kDa	91.8 kDa
pI	6.85	5.4
GC content	51.6 %	56.2 %
ORF2		
Size	831 amino acids	726 amino acids
Mass	93.8 kDa	83.1 kDa
pI	9.92	9.42
GC content	40.7 %	43.5 %
ORF3		
Size	248* amino acids	302 amino acids

* partial 5'UTR and ORF3

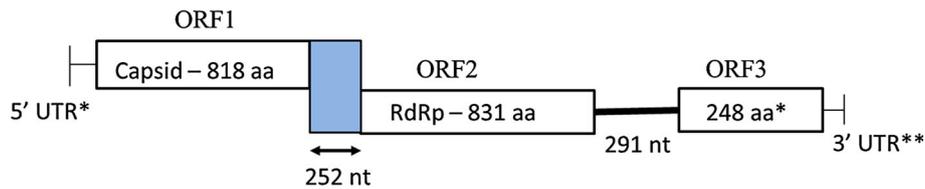


Fig. 1 Genome arrangement of golden shiner-piscine myocarditis-like virus (GS-PMCLV). The shaded are shows the overlapping region of ORF1 and ORF2. *, partial 5'UTR and ORF3; **, 3'UTR not amplified

reference PMCV sequence. Interestingly, the molecular mass of the ORF2 product (93.84 kDa) is greater than that of the reference PMCV (83.1 kDa), but their pI values are similar (9.92 and 9.42, respectively). The slippery heptanucleotide sequence was not observed in the study sequence, which is a characteristic feature of both the reference PMCV and IMNV genomes. A ribosomal -1 frameshift effect was observed in the reference PMCV sequence due to the presence of the slippery heptanucleotide, which leads to formation of a fusion protein.

A gap region (291 nt) is present between the stop codon of ORF2 and the start codon of ORF3 in the study sequence, which is 41 nt longer than the gap region in the reference PMCV sequence (250 nt). The encoded protein sequence is 248 aa long with no stop codon and hence appears to be incomplete at the 3' end. BLASTx analysis of ORF3 revealed a match with a C-X-C chemokine 3 motif, which is present at the N-terminal end from aa 17 to 60. The reference PMCV sequence also has a C-X-C motif at its N-terminal end, and this sequence showed similarity to that of chemokine 11.

Phylogenetic analysis

Based on phylogenetic analysis of ORF1, the study sequence most closely grouped together with PMCV, with a maximum aa sequence identity of 49.0 % (Fig. 2, Table 2). Similar to ORF1, the ORF2 of the study sequence grouped most closely with PMCV, with a maximum aa sequence identity of 58.5 %. After PMCV, the ORF2 study sequence had a maximum aa sequence identity of 27.5 % with Giardia lamblia virus (GLV) (Fig. 3, Table 2). The ORF3 sequence had minimal sequence similarity to the C-X-C motif of chemokine 3 and only 16.3-17.2 % aa sequence identity to reference PMCV sequences.

RT-PCR

The RT-PCR assay designed in this study targeting the RdRp gene was sensitive and specific for detection of GS-PMCLV. A single band of 921 bp confirmed the presence

of GS-PMCLV. Three samples that were positive by Illumina sequencing were confirmed using this assay, and an additional three pooled samples were positive. A total of six locations out of 56 were positive, resulting in a 10.7 % prevalence of GS-PMCLV at the surveyed Minnesota bait shops. The sequences of all RT-PCR-positive samples were 100 % identical with sequences obtained by Illumina sequencing.

Discussion

Next-generation sequencing is one of the most advanced techniques for detection and complete genome amplification of known and unknown pathogens present in a sample. We used this technique for the detection of all known and unknown pathogens circulating in apparently healthy golden shiners collected from bait shops in Minnesota. During this screening, we identified a novel PMCV related to members of the family *Totiviridae*.

The nearly complete genome of 5819 nt includes a partial 5' UTR, the complete ORF1 and ORF2, and a partial ORF3. The study sequence is 869 nt shorter than the reference PMCV. This could be because the 5' and 3' UTRs and ORF3 are incomplete, which needs to be confirmed in the future by complete sequencing of both UTRs. Interestingly, no slippery heptanucleotide was identified in ORF1 of the study sequence, despite being present in PMCV [7]. In the reference PMCV, ORF1 and ORF2 were predicted as nonoverlapping ORFs, which is also in contrast to the study sequence, in which ORF1 and ORF2 are predicted to overlap [7]. Phylogenetic analysis of ORF1 and ORF2 aa sequences showed that GS-PMCLV was most closely related to PMCV [7]. Based on the most conserved ORF2-encoded RdRp aa sequences, GS-PMCLV had a maximum identity of 58.2 % to the reference PMCV, followed by 27.5 % with GLV. The criteria established by the International Committee on Taxonomy of Viruses (ICTV) for species demarcation are not consistent for all genera in the family *Totiviridae*. However, the criteria for demarcation of a species within a genus includes less than 50 %

Fig. 2 Phylogenetic tree constructed based on 800 amino acids of ORF1 (capsid protein). The tree was constructed in MEGA 6.06 by the maximum-likelihood method, using the LG+G model and 1000 bootstrap replicates. The study sequence is highlighted in bold

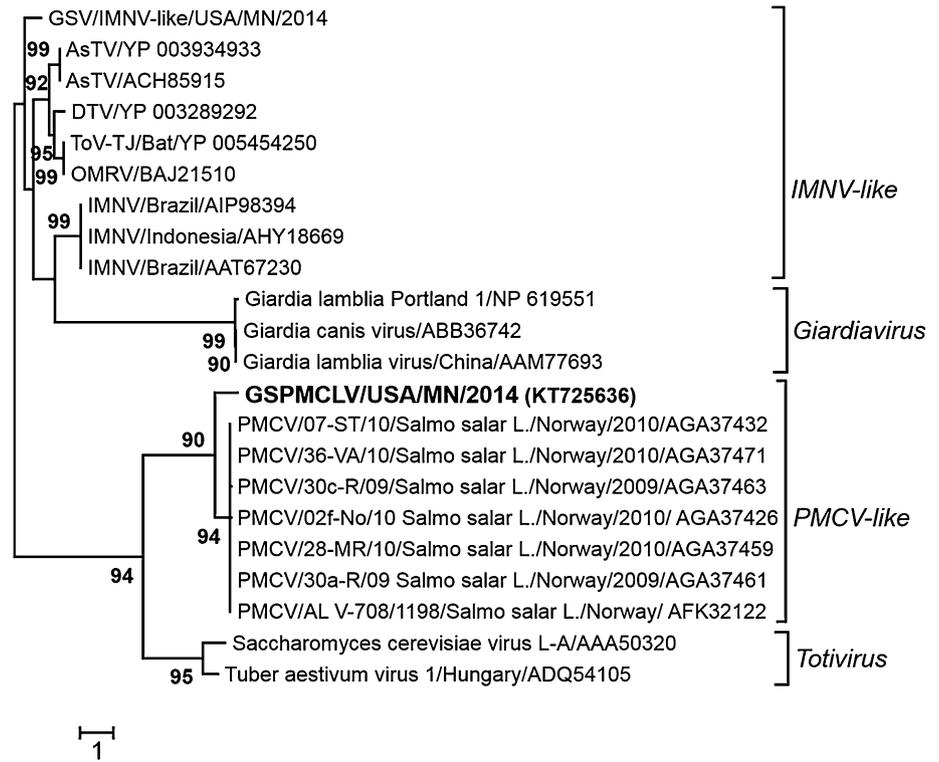


Table 2 Comparison of GS-PMCLV to related viruses of the family *Totiviridae*

Virus	Percent amino acid sequence identity		
	ORF1	ORF2	ORF3
PMCV	48.8-49.0	58.2	16.3-17.2
GLV	7.8	27.5	-
IMNV	8.3	-	-
GS-IMNLV	10.7	-	-

GS-PMCLV, golden shiner piscine myocarditis like virus; PMCV, piscine myocarditis virus; GLV, Giardia lamblia virus; IMNV, infectious myonecrosis virus; GS-IMNLV, golden shiner infectious myonecrosis-like virus

identity at the protein level [18]. If we consider the criteria for unassigned species in the family *Totiviridae*, both reference PMCV and GS-PMCLV should be grouped together in a novel genus within the family *Totiviridae*, since they share more than 50 % aa sequence identity in RdRp and have less than 50 % identity to the most closely related GLVs.

To the best of our knowledge, this is the first report of a totivirus from freshwater fish. In this study, a GS-PMCLV

prevalence of 10.7 % was observed in pooled samples of golden shiners collected at Minnesota bait shops from May 2014 to February 2015. The prevalence of GS-PMCLV within golden shiner populations at a single bait shop or source farm is unknown. Additional surveys of source populations by the specific and sensitive RT-PCR assay developed in this study are warranted to estimate the prevalence of this virus in the golden shiner industry and to study its molecular epidemiology. Understanding the distribution and monitoring the health of GS-PMCLV-positive populations would help to inform management decisions.

Although the PMCV is known to cause high mortality and morbidity in Atlantic salmon, the golden shiners in this study infected with GS-PMCLV did not show any indication of gross pathology. It is possible that GS-PMCLV does not cause disease in fish and is simply a virus of minimal concern. It is equally plausible that golden shiners are asymptomatic carriers of GS-PMCLV and the ultimate potential of this virus in other species or locations has not yet been realized. Research on seasonality, host species, stress, virus strains, and other factors affecting susceptibility and disease-causing potential are worthy of further investigation.

As molecular diagnostic techniques improve and deep sequencing becomes more common, we expect

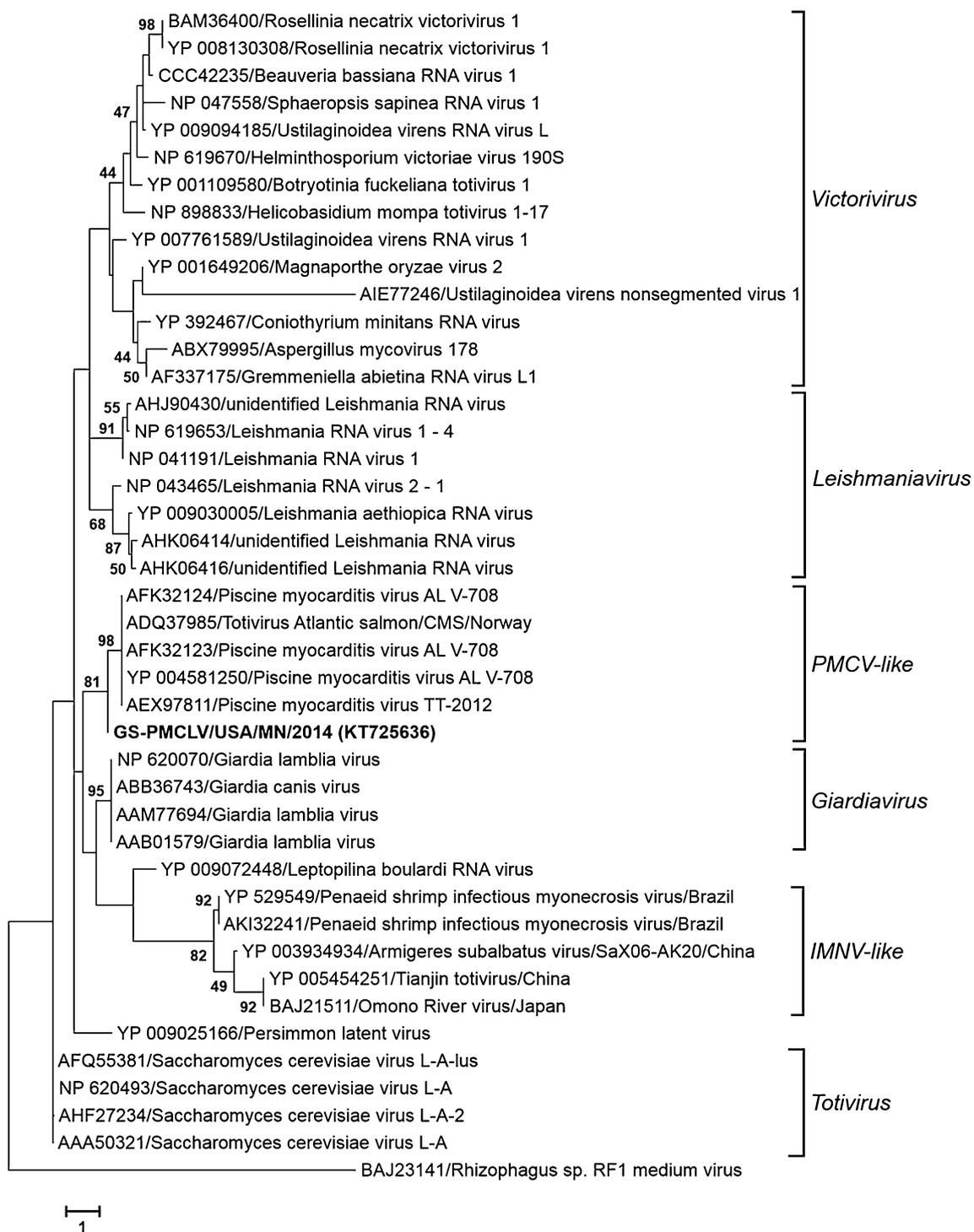


Fig. 3 Phylogenetic tree constructed based on 700 amino acids of ORF2 (RdRp gene). The tree was constructed in MEGA 6.06 by the maximum-likelihood method, using the LG+G model and 1000 bootstrap replicates. The study sequence is highlighted in bold

similar reports of novel viruses of fish to emerge. Understanding the epidemiology and potential pathology of each is critical for early detection, risk assessment, and informing evidence-based management to protect the health of wild and farm-raised fish populations.

Acknowledgments The authors thank the University of Minnesota Veterinary Diagnostic Laboratory.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Amin AB, Trasti J (1988) Endomyocarditis in Atlantic salmon in Norwegian sea farms: a case report. *Bull Eur Assoc Fish Pathol* 8:70–71
- Brocklebank J, Raverty S (2002) Sudden mortality caused by cardiac deformities following screening of preharvest farmed Atlantic salmon (*Salmo salar*) and by cardiomyopathy of postintra-peritoneally vaccinated Atlantic salmon parr in British Columbia. *Can Vet J* 43:129–130
- Bockerman I, Wiik-Nielsen CR, Sindre H, Johansen R, Tengs T (2011) Prevalence of piscine myocarditis virus (PMCV) in marine fish species. *J Fish Dis* 34:955–957
- Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Held J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2011) Geneious v5.4. <http://www.geneious.com/>
- Garseth AH, Biering E, Tengs T (2012) Piscine myocarditis virus (PMCV) in wild Atlantic salmon *Salmo salar*. *Dis Aquat Organ* 102:157–161
- Goodwin AE, Peterson JE, Meyers TR, Money DJ (2004) Transmission of exotic fish viruses: the relative risks of wild and cultured bait. *Fisheries* 29:19–23
- Haugland O, Mikalsen AB, Nilsen P, Lindmo K, Thu BJ, Eliassen TM, Roos N, Rode M, Evensen O (2011) Cardiomyopathy syndrome of Atlantic salmon (*Salmo salar* L.) is caused by a double-stranded RNA virus of the *Totiviridae* family. *J Virol* 85:5275–5286
- Isawa H, Kuwata R, Hoshino K, Tsuda Y, Sakai K, Watanabe S, Nishimura M, Satho T, Kataoka M, Nagata N, Hasegawa H, Bando H, Yano K, Sasaki T, Kobayashi M, Mizutani T, Sawabe K (2011) Identification and molecular characterization of a new nonsegmented double-stranded RNA virus isolated from *Culex* mosquitoes in Japan. *Virus Res* 155:147–155
- Mor SK, Phelps NBD (2016) Molecular detection of a novel totivirus from golden shiner (*Notemigonus crysoleucas*) baitfish in the USA. ARVI-D-16-00039R1 (**manuscript in preparation**)
- Naim S, Brown JK, Nibert ML (2014) Genetic diversification of penaeid shrimp infectious myonecrosis virus between Indonesia and Brazil. *Virus Res* 189:97–105
- Phelps NBD, Mor SK, Armien AG, Batts W, Goodwin AE, Hopper L, McCann R, Ng TFF, Puzach C, Waltzek T, Winton J, Goyal SM (2014) Isolation and molecular characterization of a novel picornavirus from baitfish in the USA. *PLoS One* 9:875–893
- Poppe TT, Seierstad SL (2003) First description of cardiomyopathy syndrome (CMS)-related lesions in wild Atlantic salmon *Salmo salar* in Norway. *Dis Aquat Organ* 56:87–88
- Rodger H, Turnbull T (2000) Cardiomyopathy syndrome in farmed Scottish salmon. *Vet Rec* 146:500–501
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Tang J, Ochoa WF, Sinkovits RS, Poulos BT, Ghabrial SA, Lightner DV, Baker TS, Nibert ML (2008) Infectious myonecrosis virus has a totivirus-like, 120-subunit capsid, but with fiber complexes at the fivefold axes. *Proc Natl Acad Sci USA* 105:17526–17531
- Tengs T, Bockerman I (2012) A strain of piscine myocarditis virus infecting Atlantic argentine, *Argentina silus* (Ascanius). *J Fish Dis* 35:545–547
- USDA (United States Department of Agriculture) (2006) Census of aquaculture (2005). 2002 Census of Agriculture: Special Studies Part 2, vol 3 AC-02-SP-2
- Wickner RB, Ghabrial SA, Nibert ML, Patterson JL, Wang CC (2012) Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses. *Totiviridae*. Academic Press, Waltham, pp 639–650
- Wiik-Nielsen J, Alarcón M, Fineid B, Rode M, Haugland Ø (2013) Genetic variation in Norwegian piscine myocarditis virus in Atlantic salmon, *Salmo salar* L. *J Fish Dis* 36:129–139
- Yang X, Zhang Y, Ge X, Yuan J, Shi Z (2012) A novel totivirus-like virus isolated from bat guano. *Arch Virol* 157:1093–1099
- Zhai Y, Attoui H, Mohd Jaafar F, Wang HQ, Cao YX, Fan SP, Sun YX, Liu LD, Mertens PP, Meng WS, Wang D, Liang G (2010) Isolation and full-length sequence analysis of *Armigeres subalbatus* totivirus, the first totivirus isolate from mosquitoes representing a proposed novel genus (*Artivirus*) of the family *Totiviridae*. *J Gen Virol* 91:2836–2845