Genomic and immunogenic changes over the history of the Viral Hemorrhagic Septicemia (VHS-IVb) fish virus (=Piscine novirhabdovirus) in the Laurentian Great Lakes

Megan D. Niner ¹*, Carol A. Stepien ²*, Bartolomeo Gorgoglione ³,⁴, Douglas W. Leaman³

¹ Genetics and Genomics Group, Department of Environmental Sciences, University of Toledo, Toledo, Ohio, 43606, USA (Stepien’s lab group relocated to ² in fall 2016)

² Genetics and Genomics Group (G3), NOAA Pacific Marine Environmental Laboratory (PMEL), 7600 Sand Point Way NE, Seattle, WA 98115, USA

³ College of Science and Mathematics, Wright State University, Dayton, Ohio, 45435, USA

⁴ Present Address: Aquatic Animal Health Laboratory, Dept. of Pathobiology and Diagnostic Investigation, CVM and Dept. of Fisheries and Wildlife, CANR - Michigan State University, East Lansing, MI 48824, USA

*Joint First Authors contributed equally to this work

#All Correspondence:
carol.stepien@noaa.gov; mcstepien@uw.edu

Running title: VHSV evolution in the Great Lakes
Abstract

Viral Hemorrhagic Septicemia Virus (VHSV) (= *Piscine novirhabdovirus*) appeared in the Laurentian Great Lakes in 2005, constituting a unique and highly virulent genogroup (IVb), which killed >32 fish species in large 2005 and 2006. Periods of apparent dormancy punctuated smaller outbreaks in 2007, 2008, and 2017. We conducted the first whole genome analysis of IVb, evaluating its evolutionary changes using 46 isolates, in reference to immunogenicity in cell culture, and the genomes of other VHS genogroups (I–IVa) and other Novirhabdoviruses. IVb isolates had 253 genomic nucleotide substitutions (2.3% of the total 11,158 nucleotides), with 85 (16.6%) being non-synonymous. The greatest number of substitutions occurred in the non-coding region (NCDS; 4.3%) followed by the Nv- (3.8%), and M- (2.8%) genes. The M-gene possessed the greatest proportions of amino acid changes (52.9%), followed by the Nv- (50.0%), G- (48.6%), N- (35.7%) and L- (23.1%) genes. Among VHS genogroups, IVa from the northeastern Pacific exhibited the fastest substitution rate (2.01x10⁻³), followed by IVb (6.64x10⁻⁵), and I/III from Europe (4.09x10⁻⁵). A 2016 gizzard shad isolate from Lake Erie was the most divergent IVb isolate (38 NT, 15.0%, 15 AA), yet exhibited reduced virulence with *in vitro* immunogenicity analyses, as did other 2016 isolates, in comparison to the first IVb isolate (2003). The 2016 isolates exhibited lower impact on innate antiviral responses, suggesting phenotypic effects. Results suggest continued sequence change and lower virulence over the history of IVb, which may facilitate its long-term persistence in fish host populations.

**Key Words:** virus evolution, Novirhabdovirus, rhabdovirus, VHSV, Viral Hemorrhagic Septicemia, antiviral response
Introduction

Viral Hemorrhagic Septicemia Virus (VHSV), a.k.a. *Piscine novirhabdovirus*, causes one of the world’s most serious viral finfish diseases, infecting >140 aquaculture and wild fish species across the northern hemisphere [1,2]. Like many other RNA viruses, VHSV has a short genome, lack of proofreading, and rapid generation time [3,4,5]. These factors may lead to its rapid evolution, facilitating adaptation to new hosts and environments [6,7,8].

VHSV suddenly appeared in North America’s Laurentian Great Lakes in 2005, causing large fish kills of >32 species in 2005-6 [9,10,11]. It was designated as the geographically and genetically distinct subgenogroup (=substrain) IVb. Prior to this, VHS was known as three genogroups (=strains) in Europe (I-III; dating to the 1930s), and genogroup IVa in the northeastern Pacific (since the 1980s) and northwestern Pacific (1990s) (summarized by [12]).

Since its emergence, VHSV IVb outbreaks have become less prevalent [13,14,15], interspersed by small, geographically restricted occurrences in 2007, 2009, 2011, and 2017 [13,16](M. Faisal, personal communication 2017; R. Getchell personal communication 2017). Yet, the virus has continued to diversify, according to sequences of selected gene regions [14]. Its diversification appears to have followed a “quasispecies” evolutionary pattern of multidirectional, “cloud-like” diversification of closely related variants [7,8,12]. In the present study, we sequenced the entire genome of available isolates to evaluate its evolutionary patterns over time and across its geographic range. We relate these to laboratory culture virulence studies.

Knowledge of evolutionary changes across VHSV-IVb’s genome and the functional roles of genes will aid understanding of the co-evolutionary “arms race” between this pathogen and its hosts.

VHSV genome and evolutionary history

VHSV has a single-stranded, negative sense RNA genome of 11,158 nucleotides (NTs), and does not recombine [17, 18]. It contains five structural genes: nucleoprotein (*N*), phosphoprotein (*P*), matrix (*M*), glycoprotein (*G*), and large protein (*L*) 5’ *N*–*P*–*M*–*G*–*Nv*–*L*’3.

VHSV belongs to the genus *Novirhabdovirus*, whose species also include *Salmonid novirhabdovirus* (IHNV), *Hirame novirhabdovirus* (HIRRV) and *Snakehead novirhabdovirus* (SHRV). All of these infect fishes, and are distinguished by their unique non-virion (*Nv*) gene. *Nv* encodes a distinctive non-structural protein [2, 18, 19], whose evolutionary history was unstudied prior to our investigation.
The functionality of the unique Nv-gene has been of scientific conjecture, which varies among the four species. Although non-essential for replication, Nv is required for pathogenicity in VHSV [19] and IHNV [20], but not in SHRV [21, 22]. Interchanging Nv between VHSV and IHNV reduced pathogenicity in both recombinants [23]. Nv appears to function in suppressing host innate immune response (IIR) [23, 24, 25] and preventing apoptosis [19], but its exact mechanisms are unknown. Single sequence changes in Nv can impact the capacity to suppress host cell responses [26, 27]. Although Nv characterizes Novirhabdoviruses, its sequence is not conserved among species [18], and our literature search found no prior investigations of its potential evolutionary origin; which is tested here.

VHSV comprises four geographically and genetically distinct lineages: I–III in Europe and IV in North America and Asia (Figures 1 and 2) [12, 28, 29]. VHSV-I is the oldest-discerned, possessing the largest known geographic range, the most named phylogenetic subgroups, and greatest number of described host species [12, 15, 18]. VHSV-I infections have caused significant economic losses to European aquaculture [30, 31].

Phylogenetic analyses found that I, II, and III shared common ancestry, separate from genogroup IV [12, 14, 15]. VHSV-IV contains three described genogroups, with IVa initially described from the northeastern Pacific coastal region in the 1980s [32, 33, 34]. In the late 1990s, IVa also appeared along the Asian Pacific coast, off Korea and Japan [35]. VHSV-IVc occurs in northwestern Atlantic marine waters, dating to its discovery in 2000 [36]. Phylogenetic analyses have placed IVc as the sister group to IVb, which together (IVb–c) then comprises the sister group to IVa [12, 14, 15]. An unclassified potential fourth genogroup, appearing derived from IVa, was isolated from wild lumpfish (Cyclopterus lumpus) in Breidafjördur Bay, Iceland [37].

**VHSV infection and host immune responses**

VHSV infection is characterized by petechial hemorrhaging diffuse to external and internal organs, with surviving fishes shedding virus up to three months post-infection [38]. Pathology studies largely have used the original VHSV-IVb 2003 isolate (C03MU) [39, 40, 41] or experimentally mutated versions [19, 42, 43], leaving uncertainty about the traits of new variants. Just two investigations have compared the original isolate to adaptive traits of other naturally occurring IVb isolates; Imanse et al. [44] found reduced virulence and lower mortality in a 2010 Lake Ontario isolate, and Getchell et al. [45] discerned lower viral titers with no mortality differences in three other round goby isolates.
After gaining entry to the host cell, the virus must successfully inhibit its innate immune defense in order to replicate. Cellular recognition of viral byproducts triggers several host signaling cascades, especially the interferon (IFN) pathway [46]. An array of Type I IFNs are highly induced by VHSV infection, triggering pathogenesis [47]. IFNs bind to specific surface receptors of effector cells, initiating transcription of IFN-stimulated genes to initiate antiviral response [48]. Thus, comparatively measuring IFN production provides insights on how effectively the virus slows down the host response, serving as a proxy of potential host-pathogen evolution.

**The Red Queen evolutionary “arms race”**

Host-pathogen relationships often are termed an evolutionary “arms-race” described by Van Valen’s [49, 50] “Red Queen” hypothesis, in which selection pressures from pathogens lead either to death or adaptation in host populations, and *vice-versa*. Since RNA viruses mutate just below the maximum limit for functionality, they readily can adapt to host defense modulations [51]. VHSV-IVb has been mutating over time [11, 14, 15], however, little is known about how its more recent isolates behave during infection. The rising importance of the aquaculture industry has raised concerns about fish pathogens from economic and sustainability perspectives [52, 53].

Biotechnology developments, including high-throughput sequencing (HTS), now allow analyses of entire viral genomes to investigate evolutionary patterns in pathogen lineages [54]. Our objective here is to understand evolutionary relationships and trajectories across the VHSV-IVb genome, and evaluate spatial and temporal patterns in relation to host-pathogen coevolution. Only five VHSV-IVb genomes were previously fully sequenced [26, 45, 55], and just one other publication evaluated similarities and differences based on those few isolates [45]. Ours is the first study to compare phylogenetic trends among entire VHSV-IVb genomes, analyzing isolates across the Great Lakes throughout their history and geographic range, and evolutionary patterns among genes. We compare the evolutionary patterns of IVb variants to other VHSV genogroups and Novirhabdoviruses. We additionally examine pathology changes of 2016 VHSV-IVb isolates to the initial 2003 isolate, in relation to IFN induction and virulence.

**Materials and methods**


Sampling and nomenclature

Samples from 2015–17 were obtained by from Niner and Stepieen [see 56]. Historical VHSV-IVb isolates were provided by G. Kurath (USGS, Seattle, WA) as frozen infected media from Bluegill Fry (BF-2) cell culture. Additional complete VHSV (I–IVa; Table 1) and Novirhabdovirus (Table 2) genomes were downloaded from GenBank and aligned using MEGA X. VHSV-IVb isolates are named here with unique identifiers, by first letter of lake name, the last two digits of its isolation year, followed by the first two letters of the host species’ common name. Example: the original isolate, collected from a Lake St. Clair muskellunge in 2003 here is C03MU. If more than one isolate shared the above identifiers, an additional lower-case letter was added (Table 1). We defined a haplotype as “a unique gene sequence differing by one or more nucleotide substitutions” from C03MU.

Virus isolation

RNA was extracted using a previously optimized Trizol protocol [57], re-suspended in RNase-free water, and quantified with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). Frozen media was thawed on ice, and 30 μL from each sample was diluted 1:5 with serum-free MEM (ThermoFisher Scientific), and added to individual wells of a 12 well plate with confluent BF-2 monolayers. Cells were incubated 1 hr with the infected media at 20°C, then replaced with complete MEM 10% (v/v) cosmic calf serum (GE Healthcare) and 1% penicillin/streptomycin antibiotics (Invitrogen). Infected cells were incubated at 20°C for ≤one week and sampled when ≥80% CPE was achieved. Media was collected and added to a 1.5ml tube with 250 μL of versene (Fisher Scientific) for 10 min prior to 4 min 4000 rpm centrifugation at 4°C. Supernatant was discarded, and the versene/cell mixture spun again as before. Remaining supernatant was discarded again, and 250 μL Trizol® (ThermoFisher Scientific) was added to the remaining pellet.

Genome sequencing

cDNA was synthesized from total RNA extracted from tissue samples using SuperScript IV (Invitrogen), following manufacturer’s instructions. Genomic cDNA was amplified in four segments using primers from Schönherz [58], substituting VHSV_Frag1I_nt18_+s with a more specific primer (5’GAGAGCTGCAGCACTTCACCG C3’), and 1μl cDNA in 25μL reactions with One Taq DNA polymerase (New England Biolabs). Amplicons were examined under UV
light on 1% agarose gels stained with ethidium bromide. Target PCR products were gel excised and purified using QIAquick Gel Extraction kits (Qiagen).

Additional PCRs amplified the front 700 nucleotides (NTs) and end 400 NTs of the genome, with 45 s extension time. The front segment utilized VHSV_Frag1I_nt18_+s (5’GAGTTATGTACRCAGGAGG3’) [58] and anti-sense 5’TGACCGAGATGGCAGTA3’, and end primers were designed based on the VHSV-IVb genome (GenBank: GQ385941) (End sense: 5’CCCAGATGCTATCACCAGAA3’, End anti-sense 5’ACAAAGAATCCGAGGAGG3’). Cleaned products were Sanger sequenced at Cornell DNA Services (Ithaca, NY) and sequences aligned and analyzed by us using MEGA X [59].

Genomic sequencing was outsourced to Ohio State University’s Molecular and Cellular Imaging Center (Wooster, OH). Sequences were uploaded by us to the Galaxy web platform and analyzed with usegalaxy.org programs [60]. Segments were aligned to the reference VHSV-IVb genome (C03MU, GenBank: GQ385941) using MAP WITH BWA-MEM [61]. For each of the 46 isolates, consensus sequences were generated followed by manual checking of each single nucleotide polymorphism (SNP) and coverage read using Integrative Genomics Viewer (IGV) [62, 63]. Consensus sequences, front, and end segments were concatenated, aligned, and trimmed in MEGA X [59].

**Genetic analyses**

The Basic Local Alignment Search Tool (BLAST) was used to detect similar sequences to Nv, excluding the *Novirhabdovirus* genus. JMODELTEST v3.7 [64] employed the Akaike Information Criterion (AIC) to determine best-fit evolutionary models [65]. Phylogenetic analyses evaluated the most parsimonious evolutionary relationships (per [12, 14]) with Maximum Likelihood (ML) (PHYML v3.0) [66] and Bayesian (MRBAYES v3.1) [67] analyses. For the latter, Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) analyses were run for five million generations, sampling every 100 to obtain posterior probability values. Burn-in was determined by plotting log likelihood values to identify stationarity, discarding the first 25%. Branch support for ML was calculated from non-parametric bootstrapping replications, which were limited to 500 for the respective *Novirhabdovirus*, 1450 for IVb, 2000 for Nv, and 500 for the Nv amino acid (AA) trees.
Haplotype networks were analyzed in POPART (www.popart.otago.ac.nz) using TCS [68]. We discerned whether samples significantly diverged over time or space using pairwise $\theta_{ST}$ ($F_{ST}$ analogue) [69] comparisons in ARLEQUIN [70], reporting values prior and after sequential Bonferroni corrections [71].

**Evaluating evolution and selection**

Comparative divergence times were estimated using BEAST v1.10.4 [72], with JMODELTEST output, a relaxed molecular clock with lognormal distribution, sampled every 50,000 of 500,000,000 generations. Outputs were assessed with TRACER v1.5 (in BEAST) to ensure stationarity. Collection dates were used as calibration points and tree branches set following PHYML output. Numbers of nucleotide substitutions per site per year ($k =$ substitutions site–1 yr–1) were determined from pairwise (p) distances. Nucleotide (NT) and amino acid (AA) substitutions were evaluated for all IVb isolates and compared to C03MU. IVa isolates were compared to KRRV9601 (GenBank: AB179621) and I/III was compared to the European Ia isolate Heddam (GenBank: Z93412).

Two codon-based methods examined the possibility of selection pressures for each gene. Fast, unconstrained Bayesian approximation (FUBAR) [73] identified positive or purifying selection; pressures selecting for beneficial traits. However, FUBAR’s assumption of constant selection may not accurately represent IVb, since selection pressures may differ across hosts and environmental factors. To remedy this, we used MEME (mixed effects model of evolution) [73], which can detect positive selection, under strong purifying selection or the removal of detrimental variants. FUBAR and MEME were run with HyPHY on DataMonkey (www.datamonkey.org), with significance evaluated using posterior probability $>0.95$ for FUBAR and $p<0.05$ for MEME.

Amino acid sequences were submitted to the Phyre2 web portal (https://www.sbg.bio.ic.ac.uk/phyre2/) for G-gene protein modeling, prediction, and analysis [74]. For open reading frame (ORF) analysis, full length cDNA nucleotide sequences per isolate were submitted to the NCBI ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/) to search for alternate products $>300$nt in length. Sequences returned coding in reverse were discounted since VHSV is single-stranded.

**Cell lines and cell culture experiments**
Epithelioma papulosum cyprinid (EPC) (ATCC: CRL-2872) and BF-2 (ATCC: CCL-91) cells were cultured as above. Three viable independent viral stocks of 2016 isolates, E16GSa (Cell16a) and E16LB (Cell16b and Cell16c) were derived from pooled organs (kidney, liver, and spleen) of sampled fish. Cell culture amplified stocks are denoted as “Cell”. Due to low concentration, the original E16LB sample was just partially sequenced, matching E16GSa within those regions [75]. The cell culture amplified reference control (CellC03) was derived from C03MU. VHSV-IVb isolates were amplified for subsequent purification by infecting confluent monolayers of BF-2 cells in 15 cm tissue culture dishes with 1:1000 (v/v) dilution of un-purified virus stock in serum-free EMEM. Viral adsorption occurred for 1 h, then medium was replaced with complete EMEM. Plates were incubated at 20°C, until >75% CPE occurred, ~72 h post infection (hpi). Virus containing media and attached cells was collected and subjected to one freeze-thaw cycle before 30 min 4,000 rpm centrifugal removal of debris at 4°C. Supernatant from each isolate was clarified using 0.22 µm syringe-tip filter and viral particles purified through a 25% (w/v) sucrose pad upon 3 hr 25,000 rpm ultra-centrifugation at 4°C. Virus-containing pellets were re-suspended overnight at 4°C in PBS. Virus stocks were titered by 1:10 serial dilution using confluent EPC cells, divided into 100 µL aliquots, and stored at -80°C. Virus-induced CPE were quantified with a sulforhodamine B (SRB) assay [43]. Infected EPC cells were sampled respectively at 0, 18, 48, 72, or 96 hours post infection (hpi) at a multiplicity of infection (MOI)=1.0 with C03MU or Cell16-a, -b, and -c, and stored at -80°C. A viral yield assay compared the replication ability of the three test isolates of E16GSa to C03MU virus, by titering media from each time point in 1:10 serial dilutions on BF-2 cells. Plaques were counted 96 hpi and final viral concentration in plaque forming units per mL (pfu/mL) calculated per time point. Antiviral assays followed Ke et al. [43]. UV-irradiated media from each harvested time points was added to EPC cells in 1:3 serial dilutions for 24 h. Cells then were challenged with sucrose-purified C03MU VHSV for 96 h, fixed, and stained with crystal violet (Sigma-Aldrich). CPE plaques were counted and normalized to counts obtained from untreated and uninfected wells. One unit of IFN was defined as the dilution that conferred 50% protection from viral CPE, which value was used to measure the IFN activity (mL) in the testing culture media per time point.

Total RNA was isolated from infected EPC cells following the TRIzol protocol of Gorgoglione [57]. One µg of each RNA sample was reverse transcribed into cDNA by 10 min
incubation at 70°C with 100 ng of random hexamer primer (Thermo-Fisher Scientific) and water, for 7 μL total volume. Reactions were cooled to 4°C before adding 13 μL of Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV-RT) mixture [10X First Stand buffer, 10 mM dNTPs, 0.05 mM random hexamers, 25 U/μL RNasin Plus (Promega), and 200 U/μL M-MLV (Promega)], incubated at 42°C for 1 hr. cDNA was diluted 10 fold in water and archived at -80°C. 1 μL of each cDNA was tested using RT-qPCR, added to 5 μL Radiant Green Lo-ROX 2× qPCR kit (Alkali Scientific), 50 ng of each oligonucleotide, and water to total 10 μL. Primer were: VHSV-Nse/as, EPC IFN se/as, and β-actin se/as [43]. Reactions and data collection were performed on a C1000 Real Time Thermocycler (Bio-Rad), with initial 3 min 95°C denaturation, 40 cycles of 15 s at 95°C, and 30 s elongation at 60°C. Readings were recorded at the end of each elongation cycle, and threshold values obtained from an automated single point threshold within the log-linear range. Detection of VHSV-N and EPC IFN was normalized to EPC β-actin detection, and normalized to the gene expression of uninfected cell samples. Relative gene expression used the 2-ΔΔCT method [76].

Results

Genomic and genic changes

Our analyses of 43 VHSV-IVb whole genome sequences (11,083 NT), plus an additional five from GenBank (GQ385941, KY359355–57), showed no insertions or deletions (Table 1). Of these, 39 sequences were unique (0.81), differing ≥1 NT from C03MU (GenBank: MK782981–MK783014). Isolates C06NP, C06RB, C06SR, C06YP, C06FD, M08AM, C08LEa, C08LEb, and C09MU were identical, designated here as the “C06NP group” (MK782990). No new open reading frames (ORFs) were detected. G-gene structures all were identified as glycoproteins.

The 39 IVb gene sequences contained 253 single nucleotide polymorphisms (SNPs) (Table 3), with 85 (0.36) being nonsynonymous, encoding different amino acids (AA). Most SNPs and AA changes occurred in L (112 SNPs, 38 AA) and the least Nv (14 SNPs, 7 AA), as anticipated from their respective lengths (L: 5955 NTs, Nv: 369). The non-coding region (NCDS), which does not encode amino acids, contained the highest proportion of SNPs (32 SNPs, 0.043 of the region), whereas P and L had the lowest (0.019). Nv encoded the highest proportion of AA changes (0.057) and P the least (0.014). The highest proportion of
nonsynonymous: synonymous changes (dN/dS) occurred in \( M \) (0.53) and the lowest in \( P \) (0.23). The overall dN/dS for the complete sequence was 0.166.

In comparison, subgenogroup IVa sequences (Table 3B) possessed 787 SNPs; 203 were nonsynonymous (0.26). Genogroup I/III (Table 3C) contained 1292 SNPs, with 225 AA changes (0.17). Most substitutions for both were in \( L \) (IVa: 424 SNPs, 119 AAs; I/III: 618 SNPs, 70 AAs). For IVa, \( Nv \) had the fewest SNPs (31 SNPs, 10 AAs), and \( M \) the fewest for I/III (65 SNPs, 15 AAs). In IVa, the largest proportion of SNPs occurred in the NCDS, and \( Nv \) and \( P \) had the most AA changes (0.081). For I/III, \( Nv \) contained the largest proportions of SNPs (0.206) and AA changes (0.220). The lowest SNP proportions occurred in \( N \) (0.044) for IVa, and \( P \) (0.019) for I/III. The least AA changes were in IVa-\( N \) (0.042) and I/III-\( L \) (0.035). For both genogroups, the \( P \) had the highest dN/dS ratio (IVa: 0.400, I/III: 0.411) and \( L \) the lowest (IVa: 0.281, I/III: 0.113).

IVb possessed the fewest transversions (Tv: IVb=42, IVa=131, I/III=254), tying with IVa in overall ratio of transversions: transitions (Tv/Ts=0.166). I/III had a larger Tv/Ts (transitions) ratio (0.197). Most Tv occurred in \( L \) for IVb (\( N\)=18, Tv/Ts=0.161), IVa (\( N\)=65, Tv/Ts=0.153), and I/III (\( N\)=106, Tv/Ts=0.172). All three groups differed in which genomic regions had the fewest Tv (IVb: \( P \) and \( M \), \( N\)=2, Tv/Ts=0.118; IVa: \( Nv \), \( N\)=3, Tv/Ts=0.097; I/III: \( M \), \( N\)=10, \( N\)=2, Tv/Ts=0.154). The largest proportion of Tv/Ts occurred in IVb-\( Nv \) (Tv=4, Tv/Ts=0.286), and NCDS for both IVa (\( N\)=0.273, Tv/Ts=0.273) and I/III (\( N\)=36, Tv/Ts=0.310). The lowest Tv/Ts ratios were: IVb-NCDS (Tv=3, Tv/Ts=0.094), IVa-\( Nv \) (Tv=3, Tv/Ts=0.097), and I/III-\( M \) (\( N\)=10, Tv/Ts=0.154).

Per isolate changes were examined for recent IVb isolates. The average number of SNPs/isolate was 16.1, with 6.1 (0.38) AA changes (Sup Table A). The C06NP group possessed the fewest SNPs, with a single non-synonymous change in \( L \). Isolates from 2012–2016 contained the highest SNP proportions (averaging 27.9 SNPs, 9.5 AAs), ranging from 14 (M16RGa, 4 AAs) to 38 (E16GSa, 13 AAs). Isolates recovered from cell culture differed by a few SNPs from their original sequences; CellC03 diverged by four SNPs from C03MU and Cell16a–c were two, six, and eight differences from E16GSa, respectively.

SNP changes are shown in Figure 3. C03MU is centrally located nearby a large cluster of isolates, including the C06NP group. Sequences recovered from 2006–08 Lake Erie outbreaks cluster together, differing by 1–4 SNPs. E12FD radiates from this cluster by 31 additional SNPs.
Both Budd Lake isolates (B07BS, PS) form an individual branch, sharing one synonymous SNP in *G*. The Lake Ontario isolates (O06SB, O13GS0) [45] group together, sharing 10 SNPs before diverging by 31 SNPs. Five shared changes were nonsynonymous: one in both *M* and *L* and three in *G*. Lake Michigan isolates appear scattered throughout the network, but remain closer to C03MU, with the 2016 isolates differing by 15–16 NT (eight shared). E16GSa–e and derived cell culture isolates form a distant, smaller cluster, sharing 14 NT changes before E16GSc diverges. The remaining seven isolates shared 19 NT changes before differentiating, with E16Gsb as central.

SNPs distinguished isolates of IVa (Sup. Figure A) and I/III (Sup. Figure B). For the IVa network, the two Japanese isolates were very distant (KRRV9822: 84nt, JFEh001: 253nt) from the nearest Korean isolate (KJ2008). Two identical isolates occurred in I/III, recovered one year apart in Norway (BV06048–52, FA281107). Seven Ib Swedish isolates from 1998–2000 and a 1996 Japanese fish farm isolate form a cluster, differing by 2–22 NT.

The fastest overall evolutionary rate characterized IVa (2.01x10^-3 substitutions/site/year), followed by IVb (6.64x10^-5) and I/III (4.09x10^-5) (Table 3). For IVa, NCDS evolved with the highest rate (7.79x10^-3), followed by *P* (2.32x10^-3) and *Nv* (2.03x10^-3). For IVb, NCDS appeared faster (1.40x10^-4), followed by *Nv* (9.76x10^-5) and *G* (8.51x10^-5). The order for I/III was: *Nv* (7.78x10^-5), *M* (5.45x10^-5), and *N* (4.59x10^-5). The slowest regions were IVa: *G* (1.02x10^-3), *M* (7.68x10^-4) and *N* (5.97x10^-4), IVb: *P* (7.18x10^-5), *N* (5.83x10^-5), and *L* (5.02x10^-5), and I/III: *P* (4.07x10^-5), NCDS (3.60x10^-5), and *L* (3.58x10^-5).

Selection pressures for each CDS were examined (Table 4) for IVb, IVa, and I/III. Purifying selection characterized IVb’s *N* (codon 313), *G* (codon 342), and *L* (six codons: 8, 119, 333, 460, 1284, and 1758) genes. For IVa and I/III, FUBAR implicated purifying selection for all genes (Table 4B, C). In IVb, one codon (*L*, 1758) indicated purifying selection and three codons implied diversifying selection – *G* (103, 431) and *Nv* (25). A IVa codon reflected diversifying selection (*G*, 12), along with one in I/III (*N*, 46). MEME implicated diversifying selection for IVb-G431 and I/III-G477, and three different *L* sites each in IVa (147, 593, 1154) and I/III (112, 474, 1012); none matched.

**Evolutionary relationships**

Significant genetic differentiation characterized IVb over time, indicted by pairwise genetic divergences (Table 5A). The later time group (2012–16) significantly diverged from
early (2003–06) and middle (2007–11) groups. Greater divergence was found between the early and later time groups, increasing over time. Population region groupings (Lakes Michigan/Budd, St. Clair, and Erie/Ontario) significantly diverged from each other (Table 5B), with greatest difference between Lakes Michigan/Budd and Erie/Ontario. Significant regional geographic differences occurred between Lakes Erie/Ontario versus other regions for M and NCDS, whereas Lakes Michigan/Budd significantly diverged from others with L, and G differed for Lake St. Clair. N differed between Lakes Michigan/Budd and St. Clair.

The Novirhabdovirus phylogenetic tree (Figure 4) defines each species with 1.00 posterior probability (pp) and 100% bootstrap support (bs). IHNV and HRRV are sister species, comprising the sister group to VHSV+SHRV. The 79 genomic VHSV sequences form two clades: European genogroups I–III and North American/Asian IV. Within I–III, the single VHSV-II isolate is basally located. Two of the five VHSV-III genotypes group with two 1a isolates (Hededa DK35928), thus not monophyletic (i.e., not a clade). In contrast, the seven 1b isolates are a well-defined clade, and the III isolates comprise another. Genogroup IV (Figure 3) possesses 1.00 pp and 100% bs support, with IVa (N=25) and IVb (N=39) on separate branches, having 1.00 pp and 100% bs support.

The IVb phylogenetic tree (Figure 5) contains several inner clades, with M08RB appearing basal (weakly supported), followed by O06SB, CellC03, and O13GS, respectively (0.60–90 pp/<50% bs). Two Lake Erie isolates from 2014–15 (E14GS, E15RG) group together (1.00 pp/74% bs), followed by two Lake Michigan 2007–08 isolates, located prior to the main cluster of the remaining 39 isolates, including C03MU. Within the major group are two clades containing two isolates (B07BG/B07PS: 0.80–.89 pp/66% bs; M11YP/C06GS: 0.80–.89 pp/<50% bs) and two larger clades. The first larger clade incorporates the 2016 isolates (0.60–.69 pp/<50% bs), further subdivided in two, separating Lake Michigan (1.00 pp/98% bs) from Lake Erie and cell culture derivatives (1.00 pp/100% bs). A second well-defined clade (1.00 pp/83% bs) contains 14 isolates, including most early to middle (2006–08) Lake Erie genotypes, along with its very distinct 2012 isolate. The 2012 genotype is the most distant, comprising the sister taxon to E08ES (1.00/94%), which are linked by four shared SNPs and 2 AAs (M: one SNP, one AA; L: two SNPs, one AA; NCDS: one SNP). E12FD and E08ES share an additional five SNPs: two in N (two AA), and two in L (one AA).
BLAST searches returned no similar sequences to Nv when novirhabdoviruses were omitted. The NT phylogeny (Figure 6A) resolved all genes as comprising distinct clades, except for Nv. The phylogeny contained two groups of clades: M+G and N+P+L. SHRV-Nv grouped with G. IHNV-Nv and HIRRV-Nv are sister groups, sharing a common ancestor with the N clade. VHSV’s Nv gene appeared a divergent part of the L clade. All Nv clades were fully supported by Bayesian analysis.

When AAs alone were considered (Figure 6B), the IHNV and HIRRV-Nv clade was sister to VHSV-Nv and the G clade. SHRV-Nv formed its own clade, more closely related to the P clade, but was not well supported.

**Differences in cytopathicity and immune response**

A series of cell-based studies were used to determine whether sequence variations impacted viral function: SRB assays to assess cytopathicity, virus yield assays to estimate viral production, antiviral assays to assess IFN production, and qPCR to determine levels of IFN and VHSV-IVb mRNA production. Sequencing confirmed each isolate’s identity and revealed changes acquired during cell culture propagation. The cell culture amplified control, CellC03, differed from C03MU by four NTs and two AAs, M (2312, NT:C→T, AA:T→I) and G (3397, NT:G→A, AA:K→D; 4007, NT:C→G, AA:G→G; 4394 NT:G→A, AA:V→V). Cell16a–c differed by 34–38 SNPs (13–14aa) from C03MU (Sup Table A). From E16GSa, Cell16a differed by two NTs: N (737, NT:C→A, AA:T→T) and G (4117, NT:G→A AA:D→N), Cell16b differed by four additional changes (six total) in L (6962, NT:C→A, AA:A→E; 7038, NT:G→A, AA:E→E; 7047, NT:T→C, AA:C→C; 7647, NT:T→C, AA:H→H), and Cell16c had one further NT change (one total) in L (6456, NT:C→T, AA:L→L).

SRB staining examined CPE elicited at different MOIs across the four isolates. The 2016 isolates exhibited less CPE than CellC03 when tested at MOIs >1x10^-3 (Figure 7). Although not statistically different across all isolates and at all dilutions, the 2016 viral isolates tested were generally less cytotoxic than CellC03. Consistent with this observation, viral yield assays demonstrated that at 96 hpi, CellC03 produced significantly more virus than the 2016 isolates (Figure 8), yielding nearly 100-fold more infectious particles as compared to Cell16a–c.

**IFN production and expression of VHSV and IFN mRNAs**

IFN production was measured in an antiviral bioassay, which revealed earlier onset of antiviral immune response in cells infected with Cell16a, yielding significantly more IFN at 24
hpi than CellC03 (Figure 9). All isolates produced comparable IFN activities after 24 hpi (Figure 9). Since IFN antiviral bioassays are relatively insensitive to minimal changes, we extended the previous studies by assessing IFN mRNA synthesis in infected cells with RT-qPCR. IFN mRNA was expressed significantly more in Cell16a–c, compared with the CellC03 control across all time points beyond 24 hpi (Figure 10). These data suggest significant difference in ability of the 2016 isolates to induce cellular IFN expression, or significant reduction in their suppression of IFN expression. Consequently, by 48 hpi, CellC03 produced the highest overall amount of viral RNA at 48 hpi, which continued until 96 hpi, by which point cell death in all cultures led to reduced viral expression. All 2016 isolates produced similar amounts of viral RNA.

**Discussion**

**Evolutionary patterns**

We analyzed 46 sequences of VHSV-IVb for the largest Novirhabdovirus genomic investigation. The greatest proportion of SNPs occurred in the NCDS (4.3%) and the Nv-gene (3.8%). In IVb-Nv, none of the isolates we sampled had >two SNPs, suggesting that sequence conservation is important. An experimental study on Nv function indicated that mutations at codon positions 36, 39, and 41 exhibited greater host immune suppression [26]; our sequences lacked SNPs at those locations. Since Nv is not essential for viral replication [19], its SNP variations may be more tolerable for the virus. Large proportions of SNPs likewise characterized NCDS in VHS-IVa and I/III, illustrating sequence flexibility. There is no evidence that NCDS regions are transcribed or play a functional role in any rhabdovirus [2], thus these SNPs likely are selectively neutral.

Diversifying selection was indicated for two IVb genes: G and Nv. Unlike our findings, a study of just four IVb genome sequences (C03MU, O06SB, O13GS, and E14GS) by Getchell et al. [45] did not uncover diversifying selection for either G or Nv. Getchell et al. [45] discerned variation in two N-gene codons and one codon each for M and L, which were unsupported in our investigation. We found evidence of purifying selection acting on single N- and G-gene codons, and on six codons in L. In our IV analyses. Thus, fewer codons were under purifying or diversifying selection than in Getchell et al.’s [45] findings. None matched between the two analyses, with ours being the larger, more robust dataset.
Purifying selection was the major evolutionary force discerned across all VHSV genogroups, subgenogroups, and samples. Our study identified ≥ one codon per gene under negative selection, for IVa and I/III. No NT sites under selection were in common between IVb and the other genogroups, with one exception: codon 1758 in L for IVb and I/III. Our dN/dS ratio further indicated purifying selection, with values being >0.5 across all genes and VHSV genogroups. Other investigations found that purifying selection regulates VHSV evolution [30, 77], as well as in other rhabdoviruses [78] and RNA viruses in general [79]. He et al. [77] compared dN/dS ratios for individual genes among VHSV genogroups, with all six genes having lower dN/dS than found with our data. However, most of the IVb (N=45) and IVa (N=18, from [80]) sequences used in our analyses were unavailable to He et al. [77].

Positive selection was indicated for changes in three G-gene codons: two in IVb and one in IVa. Positive selection on viral genes constitutes evidence that supports the host-pathogen “arms race”, which typically suppresses host immune responses [51]. Abbadi et al. [30] examined 108 VHSV-Ia full-genome G isolates, discerning positive selection on two codons (258, 259). He et al. [77] also elucidated positive selection for I at codons 258 and 476, but not for 259. Our codon results did not match either of those studies.

The host immune system exerts strong selective pressure on viral mutation rates. Host suppression by multifunctional viral proteins is a critical determinant of success for all RNA viruses [51], including Human Immunodeficiency Virus (HIV) [81]. Viral RNAs can suppress host RNA interference pathways involved in antiviral immunity [82] or block the abilities of host antiviral proteins to engage with viral dsRNAs and mount an immune response [83].

We found that overall evolutionary rates across the entire genome were slower than were previous estimates based on partial gene sequences [12]. These differences stem from our present use of more samples and complete genome sequences, which included conserved areas that were previously omitted.

Among the three VHSV genogroups, V-IVa exhibited the fastest rate – 2.01x10⁻³. He et al. [77] estimated a rate of 5.60x10⁻⁴ for 48 IVa G-gene sequences, which was slower than our IVb rate for G of 1.02x10⁻³. Our inclusion of recently published diverse sequences from Korea between 2012–16 [80] may have increased this, as those isolates differed by >20nt. Just two complete sequences of IVa were from Japan, which were even more different than the Korean isolates. This suggests that the Korean isolates are diverging from the Japanese IVa isolates, but
additional IVa genomes from Japan are needed for robust conclusion. IVa first was detected along the North American Pacific Coast [32, 33, 34], but no full-genome sequences then were available. All IVa Korean isolates were obtained from aquaculture; such transportation of infected fishes may introduce VHSV to naïve fishes and different environmental conditions, which can enhance adaptation and dramatically alter the virus, as observed in IHNV [84] and HIRRV [85].

Some Ia, Ib, and III isolates grouped together in our phylogeny, indicating that their previous phylogenetic relationships, based on partial genes [12, 30, 86], had been poorly resolved in the earlier analyses. We calculated that genomes of VHSV genogroups I/III evolved at an overall rate of $4.09 \times 10^{-5}$, more similar to that of IVb. Previous G-gene rate estimates for VHSV-I and -III [77] using fewer complete gene sequences were faster than ours ($4.54 \times 10^{-5}$), having been based on 201 I sequences ($5.77 \times 10^{-4}$) and seven III isolates ($1.63 \times 10^{-3}$). Evolution of subgenogroup Ia had been estimated at $1.74 \times 10^{-3}$ ($N=34$) [23] to $7.3 \times 10^{-4}$ from 108 G-gene Italian isolates [30]. Our combination of the European groups may have yielded our slower rate. Despite these variations, all of the above estimates are within the known range of evolutionary rates for RNA viruses [87](Duffy et al. 2008).

**Phylogenetic patterns: Novirhabdoviruses**

The phylogeny from the whole genome (Figure 3) is congruent with prior partial gene studies [12, 14]. Two main sister groups characterize the novirhabdoviruses: IHNV+HIRRV and VHSV+SHRV, with VHSV+SHRV diverging first from the common ancestor. This relationship is congruent with Kurath’s (2012 [18]) analysis of complete N-gene sequences. Two studies examined rhabdoviruses using partial L-gene sequences [88, 89], which likewise supported our sister group pairing of IHRV+HIRRV. Our phylogenetic consensus tree yielded 100% ML and Bayesian support for these relationships, whereas Kurath [18] found just 76% support for the SHRV+VHSV clade using N. Thus, our analysis of the entire genome using a larger number of isolates significantly increased confidence in resolving novirhabdovirus evolutionary relationships.

Phylogenetic results for VHSV relationships from whole genome sequences support most known genogroups and subgenogroups, except that III and Ia are not monophyletic. Prior analyses of G-gene sequences by Dale et al. [90] and Ghorani et al. [31] thus were incorrect in separating III and Ia. Genogroup Ia occurs in freshwater hosts [23], whereas both III and Ib are
marine [90]. All three are capable of infecting rainbow trout (*Oncorhynchus mykiss*), which
spends portions of its lifecycle in both environments. Our results indicate that Ib is
monophyletic, whereas III and Ia comprise a single taxon, indicating that they need to be re-
defined as a single taxon. Additional III and Ia isolates should be completely sequenced.

**Phylogenetic patterns: VHSV-IVb**

Our phylogenetic tree contains several clades of closely related IVb isolates, which
correspond to sampling location and year. We resolve an in-depth view of the virus in Lake Erie,
where a large portion of our samples were collected. The inner-most branches of the phylogeny
suggest that at least two separate infections led to the major outbreaks, since one of the 2006
isolates (E06FD) appeared more closely related to C03MU and the Lake St. Clair outbreak
samples.

The differentiated Lake Erie clade from 2006–08 and 2012 suggests its separate
evolutionary history from the 2014–15 and 2016 clades. Broad diversification characterized the
2006 outbreak, which continued along that trajectory in 2007–08, with minor variants. The lone
2012 sample, although evolutionarily distant, appears more closely related to the E06–08 clade,
than to Lake Erie samples from subsequent years. E12FD forms a sister clade with E08ES,
sharing an additional four AA changes in the coding regions of *N*, *P*, and *L*. These AA changes
possibly provided some advantage or lacked deleterious effects, since the substitutions remained
conserved for four years. Although most of that clade was recovered from Percidae hosts, several
of its isolates were in freshwater drum. In addition to the 2006 and 2008 Lake Erie cases [11],
freshwater drum die-offs were reported in 2005 from Lake Ontario [9], which appeared more
closely related to E06FD.

Both of our 2007 samples from the inland Budd Lake form a clade, located near C03MU.
Wayne County, Michigan, which is adjacent to Lake St. Clair that is a highly urbanized area with
the largest number of registered anglers in the state [91]. Anglers traveling from that area to
Budd Lake may have unknowingly transported VHSV, perhaps via live bait [12, 14, 92].

The 2014–15 Lake Erie VHSV-IVb genotypes differ from one another, but share a sister
relationship branching from the main part of the phylogeny, located closer to C03MU.
Genotypes E14GS and E15GS shared the host species of gizzard shad and round goby with the
2016 isolates, but are genetically distant from the latter. The 2016 genotypes form a clade, which
is linked by SNPs as shown in the genotype network. Interestingly, these occurred in different host species, with the E16 isolates exclusively recovered from gizzard shad, and both M16 isolates in round goby. Other studies identified potential evolutionary radiation of IVb in the round goby host, particularly in Lake Ontario [40, 75]. There have been relatively few VHS positive round goby samples detected in Lake Michigan, despite its reported deaths during the 2008 outbreaks [16]. Round goby has been implicated in spreading the virus [93], often is used as bait, and is transported among water bodies by anglers [94].

Other VHS-positive isolates were recovered from Lakes Michigan and Erie in 2016, whose viral titers were too low for genome sequencing [75], further indicating that infection levels differ spatially and temporally. One of the first 2006 outbreaks intensely affected the Lake St. Clair gizzard shad population [16], as well as its 2017 outbreak (G. Whelan, personal communication 2017). Gizzard shad populations frequently experience large die-offs from cold weather [95], water temperature fluctuations, and/or spawning activities [96], coinciding with the optimal temperature range of VHSV-IVb (12–18 °C [97]; 10–14 °C [98]. In addition to the 2016 isolates, IVb-positive gizzard shad also occurred in Lakes St. Clair, Erie, and Ontario dating back to 2006; however, only its 2016 isolates were closely related. Possible host specialization should be examined in future investigation.

All but one of our identical sequences (C06FD group) were collected from Lake St. Clair over a five-year span. This consistency suggests that the C06FD genome may have been prevalent and locally adapted to environmental conditions and host species. It differed from C03MU by a single AA in the L-gene, and was found in six different fish host species and the sole two known invertebrate host species [99, 100]. Yusuff et al. [101] discerned no effect on host specificity when swapping L-genes between C03MU and Ia’s DK-3592B. L variants appear to differ in optimal temperature range, since swapping L from IVa into IVb led to increased virus at higher temperatures, than occurred for IVa [102]. The single AA change between the C06FD group and C03ML was not shared with IVa; instead C03ML and IVa were identical at that NT, meriting future examination.

Analyses of population divergence patterns using the VHSV-IVb genome discern similar spatial and temporal patterns as found in our prior partial G-gene analyses [75]. These results showed that increased divergence characterized the later time period (2012–2016), as compared to the early (2003–2007) and middle (2008–2011) time periods of VHSV-IVb in the Great
Lakes. This pattern is evident in all individual genes as well, suggesting that the G-gene depicts overall population trends. During the later time period, VHSV-IVb has occurred more sporadically in smaller, localized outbreaks, which may reflect increased host population immunity and resistance, as predicted by the “Red Queen hypothesis” [49, 50]. Likewise, viral phage Φ2 showed greater genome divergence when exposed to changing hosts versus consistent ones, particularly in genes associated with host immune suppression [103]. On the host side, Alves et al. [104] described genomic changes in immune-related genes of wild rabbit populations following 60 years of exposure to Myoxoma Virus (MYXV).

**Evolutionary perspectives for the Nv-gene**

Our multi-gene phylogenies based on NT and AA sequences reveal little consensus regarding the evolutionary origin of the Nv gene. All results support the sister relationship between IHNV and HIRRV. VHSV-Nv appears possibly related to the L-gene in the NT sequence tree, but pairs with IHNV-Nv and SHRV-Nv in the AA tree, showing closer relationship to the G-gene. SHRV-Nv differs the most from other novirhabdoviruses, appearing more closely related to the G-gene in the NT tree and to the P-gene for the AA tree. This discrepancy might reflect differential functionality, since studies showed that Nv is not essential to pathogenicity in SHRV [20, 22]. In contrast, VHSV and IHNV require Nv for suppression of the host immune response [19]. We conclude that Nv is highly saturated and homoplastic, especially for SHRV.

**Differences in cytopathogenicity**

After 24–72 hpi, the 2016 isolates produced less VHSV-IVb RNA and induced higher IFN transcription, compared to the reference 2003 strain (CellC03) (Figure 10). Although viral yields did not significantly differ, by 96 hpi CellC03 produced more active viral particles (Figure 8) and was more cytotoxic at higher MOIs (Figure 7). The genetic changes seen between CellC03 and Cell16a–c may have affected their ability to suppress host response and cause cellular damage. These may contribute to the virus appearing less virulent over time. Cell16b–c (both isolated from E16LB) differ by a single L-gene NT and from Cell16a (isolated from ESG16a) by four L NTs (one AA). Single nonsynonymous changes can have functional consequences for VHSV. Notably, Chinchilla and Gomez-Casado [26] discerned the function of SNPs in Nv by inducing mutations in VHSV-Ia strain FR07-71. Ke et al. [43] found that when four AA M-gene mutations were introduced to C03MU, the virus was less able to suppress the...
transcription of host immune defenses. These data suggest that even subtle changes can lead to altered gene function and viral phenotypic characteristics.

It appears that the oldest VHSV-IVb isolate, C03MU, has remained the most virulent to date. More recent isolates from 2006–16 showed reduced virulence, potentially resulting from virus-host coevolution. Other studies observed similar trends. For example, Imanse et al. [44] examined vcG001 (C03MU) and vcG002 (not included in our analysis), finding faster growth but lower titers in cells exposed to vcG002. Those had similar levels of viral RNA, suggesting that vcG002 was less efficient at producing infective particles. Getchell et al. [45] also found reduced viral load in O06SB, O13GS, and E14GS. In our results, viral RNA production and virulence of Cell16a–c likewise did not significantly differ from CellC03, as evidenced by lower peak virus production in the viral yield assay. Reduced virulence over time has characterized MYXV in Australian rabbits [105], human papillomaviruses [106], and RABV in hyenas [107], leading to a longer infectious period that may aid spread of the virus. Such pattern may similarly characterize VHSV-IVb, meriting further work.

From our findings, we are unable to determine the mechanism underlying the observed reduction in virulence for the 2016 isolates. Reduced virulence might indicate less ability to suppress the host immune response, or it could be that the stronger immune response is the consequence of reduced virulence itself, allowing infected cells more time to produce antiviral factors. Further investigations into the impact of individual SNP changes seen in the 2016 isolates might answer this lingering question. Other investigations observed large impacts on virulence from single nucleotide changes in VHSV-I [27] and Vesicular Stomatitis Virus (VSV) [108]. It is possible that reduced virulence may play a role in aiding transmission.

Conclusions

As VHSV-IVb nears the end of its second decade of evolution in the Great Lakes region, it has continued to change genetically and in virulence. IVb contains fewer SNPS than have occurred in IVa or I/III, with IVa appearing to evolve the fastest, and rates of IVb and I/III being similar to each other. Our in vitro studies show that more recent isolates exhibit reduced virulence, consistent with other findings and field observations. Overall, VHSV-IVb may continue to comprise a threat to fish populations in the Great Lakes region and other waterways,
as it continues to evolve. These patterns suggest an overall trend towards persistence at relatively low levels and lesser virulence in host populations over time, in concert with predictions of the Red Queen hypothesis.

Acknowledgements

Research funding was provided by grants to CAS and DWL, including NSF-DBI-1354806 for “Gene diversity of the VHS fish virus: Evolution of cellular immune response and pathogenesis” and USDA-ARS CRIS #3655-31320-002-00D, under specific cooperative agreement #58-3655-9-748, “VHS fish virus in yellow perch aquaculture”. We thank S. Edwards for field and laboratory work, N. Marshall for computational and laboratory help, and G. Kurath and J. Winton for providing historical samples. We are grateful to F. Averick, D. Butterfield, F. Calzonetti, K. Czajkowski, A. Elz, T. Fisher, M.E. Hernandez Gonzalez, A. Izzi, M. Krishnamurthy, D. Moorhead, and S. McBride for logistical assistance. This is contribution #4974 from the NOAA Pacific Marine Environmental Laboratory (PMEL).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

CAS conceived the study, and supervised the overall evolutionary study and MDN’s graduate work; CAS and DWL wrote the NSF grant that funded the work; MDN helped fine-tune the concept, collected the data and analyses, designed the figures and tables, and wrote the first draft; DWL and BG oversaw the cell culture and immune response protocols; MDN, CAS, and DWL interpreted the data and wrote the manuscript. All authors revised and edited the manuscript. Final revision, editing, and proof-reading were provided by CAS.

Data accessibility

All sequences are deposited in NIH GenBank as MK782981–MK783014 (see Table 1). All analyses are deposited in Dryad under XXX.
Carol A. Stepień  Orcid: 0000-0002-5544-4333
References

1. Escobar LE, Escobar-Dodero J, Phelp NB. Infectious disease in fish: global risk of Viral Hemorrhagic Septicemia Virus. Reviews in Fish Biology and Fisheries. 2018; 28:637–655. https://doi.org/10.1007/s11160-018-9524-3

2. Walker PJ, Blasdel KR, Calisher CH, Dietzgen RG, Kondo H, Kurath G, et al. ICTV Virus Taxonomy Profile: Rhabdoviridae. Journal of General Virology. 2018;99:447–448. https://doi.org/10.1099/jgv.0.001020

3. Holmes E. The Evolution and Emergence of RNA Viruses. 2009; Oxford University Press, Oxford, U.K.

4. Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral mutation rates. Journal of Virology. 2010; 84:9733–9748. https://doi.org/10.1128/JVI.00694-10

5. Volz, EM, Koelle K, Bedford T. Viral phylodynamics. PLOS Computational Biology. 2013; 9:e1002947. https://doi.org/10.1371/journal.pcbi.1002947

6. Quer J, Huerta R, Novella IS, Tsimring L, Domingo E, Holland JJ. Reproducible nonlinear population dynamics and critical points during replicative competitions of RNA virus quasispecies. Journal of Molecular Biology. 1996; 264:465–471. https://doi.org/10.1006/jmbi.1996.0654

7. Lauring AS, Andino R. Quasispecies theory and the behavior of RNA viruses. PLOS Pathology. 2010; 6. https://doi.org/10.1371/journal.ppat.1001005

8. Andino R, Domingo E. Viral quasispecies. Virology. 2015; 479:46–51. https://doi.org/10.1016/j.virol.2015.03.022

9. Lumsden JS, Morrison B, Yason C, Russell S, Young K, Yazdanpanah A, et al. Mortality event in freshwater drum Aplodinotus grunniens from Lake Ontario, Canada, associated with Viral Haemorrhagic Septicemia Virus, Type IV. Diseases of Aquatic Organisms. 2007;76:99–111. https://doi.org/10.3354/dao076099

10. Groocock GH, Getchell RG, Wooster GA, Britt KL, Batts WN, Winton JR. Detection of Viral Hemorrhagic Septicemia in round gobies in New York State (USA) waters of Lake Ontario and the St. Lawrence River. Diseases of Aquatic Organisms. 2007;76:187–192. https://doi.org/10.3354/dao076187

11. Thompson TM, Batts WN, Faisal M, Bowser P, Casey JW, Phillips K, et al. Emergence of Viral Hemorrhagic Septicemia Virus in the North American Great Lakes region is associated with low viral genetic diversity. Diseases of Aquatic Organisms. 2011;96:29–43. https://doi.org/10.3354/dao02362

12. Pierce LR, Stepien CA. Evolution and biogeography of an emerging quasispecies: diversity patterns of the fish Viral Hemorrhagic Septicemia Virus (VHSV). Molecular Phylogenetics and Evolution. 2012; 63:327–341. https://doi.org/10.1016/j.ympev.2011.12.024

13. Cornwell ER, Anderson GB, Coleman D, Getchell RG, Groocock GH, Warg JV, et al. Applying multi-scale occupancy models to infer host and site occupancy of an emerging viral fish pathogen in the Great Lakes. Journal of Great Lakes Research. 2015;41:520–529. https://doi.org/10.1016/j.jglr.2015.01.002

14. Stepien CA, Pierce LR, Leaman DW, Niner MD, Shepherd, BS. Gene diversification of an emerging pathogen: a decade of mutation in a novel fish Viral Hemorrhagic Septicemia (VHS) substrain since its first appearance in the Laurentian Great Lakes. PLOS ONE, 2015;10:e0135146. https://doi.org/10.1371/journal.pone.0135146
15. Stepien CA, Leaman DW, Niner MD. Rhabdoviruses (with focus on Viral Hemorrhagic Septicemia Virus and other Novirhabdoviruses). 2020; Chapter 4 In: Woo PTK, Leong J, Buchmann K. Climate Change and Infectious Fish Diseases. CABI, England, U.K.
16. Faisal M, Shavalier M, Kim RK, Millard EV, Gunn MR, Winters AD, et al. Spread of the emerging Viral Hemorrhagic Septicemia Virus strain, genotype IVb, in Michigan, USA. Viruses. 2012;4:734–760. https://doi.org/10.3390/v4050734
17. Wolf K. Hemorrhagic septicemia virus. 1988; In: Fish Viruses and Fish Viral Diseases. (ed. Wolf K). pp 217–248. Cornell University Press, Ithaca, NY, USA.
18. Kurath G. Fish novirhabdoviruses. 2012; In: Rhabdoviruses: Molecular Taxonomy, Evolution, Genomics, Ecology, Host-vector Interactions, Cytopathology, and Control (eds. Dietzgen R G and I V Kuzmin) pp. 89–116. Caister Academic Press, Norfolk, United Kingdom.
19. Ammayappan A, Vakharia VN. Nonvirion protein of Novirhabdovirus suppresses apoptosis at the early stage of virus infection. Journal of Virology. 2011;JVI-00597. https://doi.org/10.1128/JVI.00597-11
20. Thoulouze MI, Bouguyon E, Carpentier C, Brémont M. Essential role of the Nv protein of Novirhabdovirus for pathogenicity in rainbow trout. Journal of Virology. 2004; 78(8):4098–4107. http://doi.org/10.1128/JVI.78.8.4098-4107.2004
21. Johnson MC, Simon BE, Kim CH, Leong JAC. Production of recombinant Snakehead Rhabdovirus: the Nv protein is not required for viral replication. Journal of Virology. 2000;74(5):2343–2350. http://doi.org/10.1128/JVI.74.5.2342350.2000
22. Alonso M, Kim C H, Johnson M C, Pressley, M. Leong JA. The Nv gene of Snakehead Rhabdovirus (SHRV) is not required for pathogenesis, and a heterologous glycoprotein can be incorporated into the SHRV envelope. Journal of Virology. 2004; 78(11):5875–5882. http://doi.org/10.1128/JVI.78.11.5875-5882.2004
23. Einer-Jensen K, Harmache A, Biacchesi S, Bremont M, Stegmann A, Lorenzen N. High virulence differences among phylogenetically distinct isolates of the fish rhabdovirus Viral Hemorrhagic Septicaemia Virus are not explained by variability of the surface glycoprotein G or the non-virion protein Nv. Journal of General Virology. 2014; 95(2):307–316. DOI:10.1099/vir.0.057448-0
24. Kim MS, Kim KH. Effects of NV gene knock-out recombinant Viral Hemorrhagic Septicemia Virus (VHSV) on Mx gene expression in Epithelioma papulosum cyprini (EPC) cells and olive flounder (Paralichthys olivaceus). Fish Shellfish Immunology, 2012; 32(3):459–463. https://doi.org/10.1016/j.fsi.2011.12.014
25. Kim MS, Kim KH. The role of Viral Hemorrhagic Septicaemia Virus (VHSV) Nv gene in TNF-α-and VHSV infection-mediated NF-κB activation. Fish Shellfish Immunology. 2013;34(5):1315–1319.
26. Chinchilla B, Gomez-Casado E. Identification of the functional regions of the Viral Haemorrhagic Septicaemia Virus (VHSV) Nv protein: variants that improve function. Fish Shellfish Immunology. 2017; 70:343–350. https://dx.doi.org/10.1016/j.fsi.2017.09.021
27. Baillon L, Mérour E, Cabon J, Louboutin L, Quenault H, Touzain F, et al. A single amino acid change in the non-structural Nv protein impacts the virulence phenotype of Viral Hemorrhagic Septicemia Virus in trout. Journal of General Virology. 2017; 98(6):1181–1184. http://doi.org/10.1099/jgv.0.000830
28. Meyers TR, Winton JR. Viral Hemorrhagic Septicemia Virus in North America. Annual Review of Fish Diseases. 1995;5:3–24. https://doi.org/10.1016/0959-8030(95)00002-X

29. Hedrick RP, Batt WN, Yu S, Traxler GS, Kaufman J, Winton JR. Host and geographic range extensions of the North American strain of Viral Hemorrhagic Septicemia Virus. Diseases of Aquatic Organisms. 2003;55:211–220. https://doi.org/10.3354/dao055211

30. Abbadi M, Fusaro A, Ceolin C, Casarotto, C, Quartesan R, Dalla Pozza M, et al. Molecular evolution and phylogeography of co-circulating IHNV and VHSV in Italy. Frontiers in Microbiology. 2016;7:1306. https://doi.org/10.3389/fmicb.2016.01306

31. Ghorani M, Adel M, Dadar M, Langeroudi AG, Kamyabi R, Vakharia VN, et al. Molecular evolution and phylogeography of co-circulating IHNV and VHSV in Italy. Frontiers in Microbiology. 2016;7:1306. https://doi.org/10.3389/fmicb.2016.01306

32. Brunson R, Ture K, Yancey J. VHS virus isolated at Makah national fish hatchery. Fish Health Section of the American Fisheries Society News. 1989; 17:3–4.

33. Hopper K. The isolation of VHSV from chinook salmon at Glenwood Springs, Orcas Island, Washington. Fish Health Section of the American Fisheries Society News. 1989; 17:1.

34. Meyers TR, Sullivan J, Emmenegger E, Follett J, Short S, Batt WN. Identification of Viral Hemorrhagic Septicemia Virus isolated from Pacific cod Gadus macrocephalus in Prince William Sound, Alaska, USA. Diseases of Aquatic Organisms. 1992;12:167–175. https://doi.org/10.3354/dao012167

35. Takano R, Nishizawa T, Arimoto M, Muroga K. Isolation of Viral Haemorrhagic Septicaemia Virus (VHSV) from wild Japanese flounder, Paralichthys olivaceus. Bulletin of the European Association of Fish Pathology. 2000;20:186–192. https://doi.org/10.3354/dao047087

36. Gagné N, MacKinnon AM, Boston L, Souter-B, Cook-Versloot M, Griffiths S, et al. Isolation of Viral Haemorrhagic Septicaemia Virus from mummichog stickleback, striped bass and brown trout in eastern Canada. Journal of Fish Diseases. 2007; 30:213–223. https://doi.org/10.1111/j.1365-2761.2007.00802.x

37. Guðmundsdóttir S, Vendramin N, Cuenca A, Sigurðardóttir H, Kristmundsson A, Iburg TM, et al. Outbreak of Viral Haemorrhagic Septicaemia (VHS) in lumpfish (Cyclopterus lumpus) in Iceland caused by VHS virus genotype IV. Journal of Fish Diseases. 2018;42(1), pp.47–62. https://doi.org/10.1111/jfd.12910

38. Kim RK, Faisal M. Shedding of Viral Hemorrhagic Septicemia Virus (Genotype IVb) by experimentally infected muskellunge (Esox masquinongy). Journal of Microbiology. 2012;50(2):278–284. https://doi.org/10.1007/s12275-012-1145-2

39. Kim R, Faisal M. Experimental studies confirm the wide host range of the Great Lakes Viral Haemorrhagic Septicaemia Virus genotype IVb. Journal of Fish Diseases. 2010; 33(1):83–88. https://doi.org/10.1111/j.1365–2761.2009.01093.x

40. Cornwell ER, Primus A, Wong PT, Anderson GB, Thompson TM, Kurath G, et al. Round gobies are an important part of VHSV genotype IVb ecology in the St. Lawrence River and eastern Lake Ontario. Journal of Great Lakes Research. 2014;40(4):1002–1009. https://doi.org/10.1016/j.jglr.2014.09.006

41. Coffee LL, Getchell RG, Groocock GH, Cornwell ER, Wooster GA, Lumsden JS, et al. Pathogenesis of experimental Viral Hemorrhagic Septicemia Virus IVb infection in adult sea lamprey (Petromyzon marinus). Journal of Great Lakes Research, 2017;43(3):119–126. https://doi.org/10.1016/j.jglr.2017.03.013
42. Biacchesi S, Mérour E, Chevret D, Lamoureux A, Bernard, J, Brémont M. *Nv* proteins of fish *Novirhabdovirus* recruit cellular PPM1Bb protein phosphatase and antagonize RIG-I-mediated IFN induction. International Journal of Scientific Reports 2017;7:44025. https://doi.org/10.1038/srep44025

43. Ke Q, Weaver W, Pore A, Gorgoglione B, Wildschutte JH, Xiao P, et al. Role of Viral Hemorrhagic Septicemia Virus matrix (*M*) protein in suppressing host transcription. Journal of Virology. 2017;91(19):e00279-17. https://doi.org/10.1128/JVI.00279-17

44. Imanse SM, Cornwell ER, Getchell RG, Kurath G, Bowser PR. *In vivo* and *in vitro* phenotypic differences between Great Lakes VHSV genotype IVb isolates with sequence types vcG001 and vcG002. Journal of Great Lakes Resources. 2014;40(4):879–885. https://doi.org/10.1016/j.jglr.2014.08.004

45. Getchell RG, Cornwell ER, Bogdanowicz S, Andrés J, Batts WN, Kurath G. Complete sequences of four IVb isolates and their virulence in northern pike fry. Diseases of Aquatic Organisms. 2017;126:211–227. https://doi.org/10.3354/dao03171

46. Janeway Jr, CA, Medzhitov R. Innate immune recognition. Annual Review of Immunology. 2002; 20(1):197–216. https://doi.org/10.1146/annurev.immunol.20.083001.084359

47. Zou J, Gorgoglione B, Taylor NG, Summathed T, Lee PT, Panigrahi A, et al. Salmonids have an extraordinary complex type I IFN system: characterization of the IFN locus in rainbow trout *Oncorhynchus mykiss* reveals two novel IFN subgroups. Journal of Immunology 2014;193(5):2273–2286. https://doi.org/10.4049/jimmunol.1301796

48. Collet B, Secombes CJ. The rainbow trout (*Oncorhynchus mykiss*) Mx1 promoter: structural and functional characterization. European Journal of Biochemistry. 2001; 268(6):1577–1584. http://doi.org/10.1046/j.1432-1033.2001.02021.x

49. Van Valen L. A new evolutionary law. Evolutionary Theory. 1973;1:1–30.

50. Van Valen L. Molecular evolution as predicted by natural selection. Journal of Molecular Evolution, 1973;3(2):89–101. https://doi.org/10.1007/BF01796554

51. Pereira F, Amorim A. Evolution: viruses. (eds. Maloy S, Hughes K). pp 566. 2013; Brenners Encyclopedia of Genetics. Elsevier, New York, NY, USA. https://doi.org/10.1016/B978-0-12-374984-0.00499-X

52. Lafferty KD, Harvell CD, Conrad JM, Friedman CS, Kent, ML, Kuris AM, et al. Infectious diseases affect marine fisheries and aquaculture economics. Annual Review of Marine Science. 2015;7:471–496. https://doi.org/10.1146/annurev-marine-010814-015646

53. Kennedy DA, Kurath G, Brito IL, Purcell MK, Read AF, Winton JR, et al. Potential drivers of virulence evolution in aquaculture. Evolution Applications. 2016;9(2):344–354. https://doi.org/10.1111/eva.12342

54. Bayliss SC, Verner-Jeffreys DW, Bartie KL, Aanensen DM, Sheppard, SK, Adams A, et al. The promise of whole genome pathogen sequencing for the molecular epidemiology of emerging aquaculture pathogens. Frontiers in Microbiology. 2017;8:121. https://doi.org/10.3389/fmicb.2017.00121

55. Ammayappan A, Vakharia VN. Molecular characterization of the Great Lakes Viral Hemorrhagic Septicemia virus (VHSV) isolate from USA Virology Journal. 2009; 6(1), 171. https://doi.org/10.1186/1743-422X-6-171

56. Niner, M.D. Evolutionary Patterns and Occurrences of the fish Viral Hemorrhagic Septicemia Virus in the Laurentian Great Lakes. 2019;Ph.D. Dissertation, University of Toledo, Toledo, OH.
57. Gorgoglione B. Heterogeneous Infections in Fish: Transcriptomic Studies on the Trout Immune Response to Single and Co-infections. 2014; Doctoral dissertation, University of Aberdeen, Scotland. Available at: https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.606427
58. Schönherz AA, Lorenzen N, Gulbrandsen B, Buitenhuis B, Einer-Jensen K. Ultra-deep sequencing of VHSV isolates contributes to understanding the role of viral quasispecies. Veterinary Research. 2016; 47(1):10. https://doi.org/10.1186/s13567-015-0298-5
59. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution; 2018;35(6):1547–1549. https://doi.org/10.1093/molbev/msy096
60. Afgan E, Baker D, Batut B, Van Den Beek M, Bouvier D, Čech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Research. 2018;46(1):W537–W544. https://doi.org/10.1093/nar/gky379
61. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013; arXiv:1303.3997v1 [q-bio.GN].
62. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nature Biotechnology. 2011; 29(1):24. https://doi.org/10.1038/nbt.1754
63. Thorvaldsdóttir H, Robinson JT, Mesirov JP. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics. 2013;14(2):178–192. https://doi.org/10.1093/bioinformatics/bts007
64. Posada D, Crandall K. MODELTEST: Testing the model of DNA substitution. Bioinformatics, 1998;14:817–818. https://doi.org/10.1093/bioinformatics/14.9.817
65. Posada D jModelTest: Phylogenetic Model Averaging. Molecular Biology and Evolution, 2008; 25:1253–1256. http://dx.doi.org/10.1093/molbev/msn083
66. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systems Biology. 2003;52(5):696–704. https://doi.org/10.1080/10635150390235520
67. Ronquist F, Huelsenbeck JP. MrBayes: Bayesian Phylogenetic Inference under Mixed Models. Bioinformatics. 2003;19:1572–1574. http://dx.doi.org/10.1093/bioinformatics/btg180
68. Clement M, Posada DC, Krandall KA. TCS: a computer program to estimate gene genealogies. Molecular Ecology, 2000;9(10):1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x
69. Weir B S and C C Cockerham. (1984) Estimating F-statistics for the analysis of population structure. Evolution, 38(6):1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x
70. Excoffier L, Lischer, HE. ARLEQUIN suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 2010;10(3):564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x
71. Rice WR. The sequential Bonferroni test. Evolution. 1989;43:223–225.
72. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evolution. 2018;4: vey016. https://doi.org/10.1093/ve/vey016
73. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. (2013) FUBAR: a fast, unconstrained Bayesian approximation for inferring selection. Molecular Biology and Evolution. 2013;30(5):1196–1205. https://doi.org/10.1093/molbev/mst030

74. Kelley L A, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols. 2015;10(6):845. http://doi.org/10.1038/nprot.2015.053

75. Niner MD, Stepien CA. Evolutionary trajectory of the fish Viral Hemorrhagic Septicemia virus across its history in the Laurentian Great Lakes: Spatial and temporal patterns. in review.

76. Livak J, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2–CT method. Methods, 2001;25(4):402–408. https://doi.org/10.1006/meth.2001.1262

77. He M, Yan XC, Liang Y, Sun XW, Teng CB. Evolution of the Viral Hemorrhagic Septicemia Virus: divergence, selection and origin. Molecular Phylogenetics and Evolution. 2014;77:34–40. https://doi.org/10.1016/j.ympev.2014.04.002

78. Kuzmin IV, Novella IS, Dietzgen RG, Padhi A, Rupprecht CE. (2009) The rhabdoviruses: biodiversity, phylogenetics, and evolution. Infectious Genetics and Evolution. 2009;9(4):541–553. https://doi.org/10.1016/j.meegid.2009.02.005

79. Hughes AL, Hughes MAK. More effective purifying selection on RNA viruses than in DNA viruses. Gene. 2007;404(1–2):117–125. https://doi.org/10.1016/j.gene.2007.09.013

80. Hwang JY, Lee S, Priyathilaka TT, Yang H, Kwon H, Kwon MG, Hwang SD, et al. Phylogenetic analysis and duplex RT-PCR detection of Viral Hemorrhagic Septicemia Virus in olive flounder (Paralichthys olivaceus) from Korea. Aquaculture. 2018;484:242–249. https://doi.org/10.1016/j.aquaculture.2017.11.038

81. Bush RM, Fitch WM, Bender CA, Cox NJ. Positive selection on the H3 hemagglutinin gene of Human Influenza Virus A. Molecular Biology and Evolution. 1999;16(11):1457–1465. https://doi.org/10.1093/oxfordjournals.molbev.a026057

82. Marques JT, Carthew RW. A call to arms: coevolution of animal viruses and host innate immune responses. Trends in Genetics. 2007;23(7):359–364. https://doi.org/10.1010/vir.0.1099/0.057448-0

83. Borghese F, Sorgeloos F, Cesaro T, Michiels T. The leader protein of Theiler's Virus prevents the activation of PKR by dsRNA. Journal of Virology. 2017; JVI.01010-19; DOI: 10.1128/JVI.01010-19

84. Xu L, Zhao J, Liu M, Kurath G, Breyta RB, Ren G, Yin J, Liu H and Lu T. Phylogeography and evolution of Infectious Hematopoietic Necrosis Virus in China. Molecular Phylogenetics and Evolutions. 2019;131:19–28. https://doi.org/10.1016/j.ympev.2018.10.030

85. Zhang J, Tang X, Sheng X, Xing J, Zhan W. (2017) Isolation and identification of a new strain of Hirame Rhabdovirus (HIRRV) from Japanese flounder Paralichthys olivaceus in China. Virology Journal, 2017; 14(1):73. https://doi.org/10.1186/s12985-017-0742-4

86. Nishizawa T, Kinoshita S, Kim W =S, Higashi S, Yoshimizu, M (2006) Nucleotide diversity of Japanese isolates of Infectious Hematopoietic Necrosis Virus (IHNV) based on the glycoprotein gene. Diseases of Aquatic Organisms. 2006;71(3):267–272. https://doi.org/10.3354/dao071267
87. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. Nature Reviews Genetics. 2008;9(4):267. https://doi.org/10.1038/nrg2323
88. Bourhy, H, Cowley, JA, Larrous F, Holmes EC, Walker PJ. Phylogenetic relationships among rhabdoviruses inferred using the L polymerase gene. Journal of General Virology, 2005; 86(10):2849–2858. https://doi.org/10.1099/vir.0.81128-0
89. Kitchen A, Shackelton LA, Holmes EC. Family level phylogenies reveal modes of macroevolution in RNA viruses. PNAS. 2011;108(1): 238–243. https://doi.org/10.1073/pnas.1011090108
90. Dale OB, Ørpetveit I, Lyngstad, TM, Kahns S, Skall HF, Olesen NJ, et al. Outbreak of Viral Haemorrhagic Septicaemia (VHS) in seawater-farmed rainbow trout in Norway caused by VHS Virus genotype III. Diseases of Aquatic Organisms. 2009;85(2):93–103. https://doi.org/10.3354/dao02065
91. Burkett EM, Winkler RL. Recreational Fishing in Michigan: Using an age-period-cohort approach to understand fishing participation. 2018, Michigan Technological University. Houghton MI, USA. Available at https://www.mtu.edu/greatlakes/fishery/state-level/pdf/michigan-angler-demographics-2000-2014.pdf. Accessed 20 April 2020.
92. Throckmorton E, Peters A, Brenden T, Faisal M. Direct and indirect evidence suggests continuous presence of Viral Hemorrhagic Septicemia Virus (Genotype IVb) in Budd Lake, Michigan: management implications. North American Journal of Fisheries Management, 2015; 35(3):503–511. https://dx.doi.org/10.1080/02755947.2015.1020077
93. Rude NP, Yung AJ, Whitledge GW. Movement of small-bodied fishes from Lake Michigan into Chicago area waterways: insights from otolith chemistry. Journal of Applied Ichthyology. 2017;33(6):1166–1172. https://doi.org/10.1111/jai.13499
94. Karsiotis SI, Brown JE, Pierce LR, Stepain CA. Salinity tolerance of the invasive round goby: experimental implications for seawater ballast exchange and spread to North American estuaries. Journal of Great Lakes Research. 2012;38:121–128. https://doi.org/10.1016/j.jglr.2011.12.010
95. Fetzer WW, Brooking TE, Jackson JR, Rudstam LG. (2011) Overwinter mortality of gizzard shad: evaluation of starvation and cold temperature stress. Transactions of the American Fisheries Society. 2011;140(6):1460–1471. https://doi.org/10.1080/00028487.2011.630281
96. Miller RR. Systematics and biology of the gizzard shad (Dorosoma cepedianum) and related fishes. Fisheries Bulletin, 1960;173:371–392.
97. Kane-Sutton M, Kinter B, Dennis PM, Koonce JF. Viral Hemorrhagic Septicemia Virus infection in yellow perch, Perca flavescens, in Lake Erie. Journal of Great Lakes Research. 2010;36(1):37–43. https://doi.org/10.1016/j.jglr.2009.11.004
98. Eckerlin GE, Farrell JM, Casey RN, Hope KM, Groocock GH, Bowser PR, et al. Temporal variation in prevalence of Viral Hemorrhagic Septicemia Virus type IVb among upper St. Lawrence River smallmouth bass. Transactions of the American Fisheries Society. 2011;140(3):529–536. https://doi.org/10.1080/00028487.2011.581975
99. Faisal M, Schulz CE. Detection of Viral Hemorrhagic Septicemia Virus (VHSV) from the leech Myzobdella lugubris Leidy, 1851. Parasites and Vectors. 2009; 2(1):45. https://doi.org/10.1186/1756-3305-2-45
100. Faisal M, Winters AD. Detection of Viral Hemorrhagic Septicemia Virus (VHSV) from Diporeia spp. (Pontoporeiidae, Amphipoda) in the Laurentian Great Lakes, USA Parasites and Vectors. 2011;4:2. https://doi.org/10.1186/1756-3305-4-2

101. Yusuff S, Kurath G, Kim MS, Tesfaye TM, Li J, McKenney DG, Vakharia VN. The glycoprotein, non-virion protein, and polymerase of Viral Hemorrhagic Septicemia Virus are not determinants of host-specific virulence in rainbow trout. Virology Journal. 2019; 16, 31. https://doi.org/10.1186/s12985-019-1139-3.

102. Kim SH, Yusuff S, Vakharia VN, Evensen Ø. Interchange of L polymerase protein between two strains of Viral Hemorrhagic Septicemia Virus (VHSV) genotype IV alters temperature sensitivities in vitro. Virus Research. 2015;195: 203–206. https://dx.doi.org/10.1016/j.virusres.2014.10.013

103. Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, et al. Antagonistic coevolution accelerates molecular evolution. Nature. 2010;464(7286):275. https://doi.org/10.1038/nature08798

104. Alves, J M, Carneiro M, Cheng JY, de Matos, AL, Rahman MM, Loog L, et al. Parallel adaptation of rabbit populations to Myxoma Virus. Science; 2019;363(6433):1319–1326. https://doi.org/10.1126/science.aau7285

105. Kerr PJ, Cattadori IM, Liu J, Sim DG, Dodds JW, Brooks JW, et al. Next step in the ongoing arms race between Myxoma Virus and wild rabbits in Australia is a novel disease phenotype. PNAS. 2017;114(35):9397–9402. https://doi.org/10.1073/pnas.1710336114

106. Bernard, HU Coevolution of papillomaviruses with human populations. Trends in Microbiology.1994;2(4):140–143. https://doi.org/10.1016/0966-842X(94)90602-5

107. East ML, Hofer H, Cox JH, Wulle U, Wiik H, Pitra C. Regular exposure to Rabies Virus and lack of symptomatic disease in Serengeti spotted hyenas. PNAS, 2001;98(26):15026–15031. DOI:10.1073/pnas.261411898

108. Sanjuán R, Moya A, Elena SF. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. PNAS. 2004;101(22):8396–8401. https://doi.org/10.1073/pnas.0400146101

109. Ammayappan A, Kurath G, Thompson TM, Vakharia VN. A reverse genetics system for the Great Lakes strain of Viral Hemorrhagic Septicaemia Virus: the Nv gene is required for pathogenicity. Marine Biotechnology, 2011;13(4):672–683. https://doi.org/10.1007/s10126-010-9329-4

110. Betts AM, Stone DM. Nucleotide sequence analysis of the entire coding regions of virulent and avirulent strains of Viral Haemorrhagic Septicaemia Virus. Virus Genes. 2000; 20(3), 259–262. https://doi.org/10.1023/A:100814881374

111. Campbell S, Collet B, Einer-Jensen K, Secomes CJ, Snow, M. Identifying potential virulence determinants in Viral Haemorrhagic Septicaemia Virus (VHSV) for rainbow trout. Diseases of Aquatic Organisms. 2009;86(3): 205–212.

112. Hwang JY, Kwon MG, Seo JA, Do JW, Park MA, Jung AH, et al. Differentially expressed genes after Viral Haemorrhagic Septicaemia Virus infection in olive flounder (Paralichthys olivaceus). Veterinary Microbiology. 2016;193: 72–82. https://doi.org/10.1016/j.vetmic.2016.05.024

113. Ito T, Kurita J, Sano M, Skall H F, Lorenzen N, Einer-Jensen K, et al. Typing of Viral Hemorrhagic Septicemia Virus by monoclonal antibodies. Journal of General Virology 2012; 93(12): 2546–2557. https://doi.org/10.1099/vir.0.043091-0
114. Jia P, Zheng XC, Shi XJ, Kan SF, Wang JJ, He JQ, et al. (2014) Determination of the complete genome sequence of Infectious Hematopoietic Necrosis Virus (IHNV) CH20101008 and viral molecular evolution in China. Infection, Genetics and Evolution. 2014;27:418–431. https://doi.org/10.1016/j.meegid.2014.08.013

115. Johnson MC, Maxwell JM, Loh PC, Leong JAC. Molecular characterization of the glycoproteins from two warm water rhabdoviruses: Snakehead Rhabdovirus (SHRV) and Rhabdovirus of Penaeid Shrimp (RPS)/Spring Viremia of Carp Virus (SVCV). Virus Research. 1999;64(2): 95–106. https://doi.org/10.1016/S0168-1702(99)00071-4

116. Kim DH, Oh HK, Eou JI, Seo HJ, Kim SK, Oh MJ, et al. Complete nucleotide sequence of the Hirame Rhabdovirus, a pathogen of marine fish. Virus Research. 2005;107(1):1–9. https://doi.org/10.1016/j.virusres.2004.06.004

117. Kim JO, Kim WS, Nishizawa T, Oh MJ. Complete genome sequence of Viral Hemorrhagic Septicemia Virus isolated from an olive flounder in South Korea. Genome Announcement, 2013;1(5), e00681-13.

118. Lopez-Vazquez C, Bandín I, Dopazo CP. Real-time RT-PCR for detection, identification and absolute quantification of Viral Haemorrhagic Septicaemia Virus using different types of standards. Diseases of Aquatic Organisms. 2015;114(2):99–116. https://doi.org/10.3354/dao02840

119. Morzunov SP, Winton JR, Nichol ST. The complete genome structure and phylogenetic relationship of Infectious Hematopoietic Necrosis Virus. Virus Research.1995;38(2–3):175–192. https://doi.org/10.1016/0168-1702(95)00056-V

120. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Pond SLK. Detecting individual sites subject to episodic diversifying selection. PLOS Genetics. 2012;8(7):p.e1002764. https://doi.org/10.1371/journal.pgen.1002764

121. Schütze H, Mundt E, Mettenleiter TC. Complete genomic sequence of Viral Hemorrhagic Septicemia Virus, a fish rhabdovirus. Virus Genes. 1999;19:59–65.

122. Wang C, Lian GH, Zhao LL, Wu Y, Li YJ, Tang LJ, et al. (2016). Virulence and serological studies of recombinant Infectious Hematopoietic Necrosis Virus (IHNV) in rainbow trout. Virus Research. 2016;220,19202. https://doi.org/10.1016/j.virusres.2016.04.015

123. Yingjie S, Min Z, Hong L, Zhiqin Y, Xiaocong Z, Zhe W. Analysis and characterization of the complete genomic sequence of the Chinese strain of Hirame Rhabdovirus. Journal of Fish Diseases; 2011; 34(2):167–171. https://doi.org/10.1111/j.1365-2761.2010.01218.x

124. Oh WT, Jun JW, Giri SS, Yun S, Kim HJ, Kim SG et al. Genome and Phylogenetic Analysis of Infectious Hematopoietic Necrosis Virus Strain SNU1 Isolated in Korea. Pathogens. 2019;8(4):200. Published 2019 Oct 21. doi:10.3390/pathogens8040200

125. Stone DM, Way K, Dixon PF. Nucleotide sequence of the glycoprotein gene of viral haemorrhagic septicaemia (VHS) viruses from different geographical areas: a link between VHS in farmed fish species and viruses isolated from North Sea cod (Gadus morhua L.). Journal of General Virology. 1997;78(6): https://doi.org/10.1099/0022-1317-78-6-1319

126. Ito T, Kurita J, Mori K-I, Skall HF, Lorenzen N, Vendramin N, et al. Virulence marker candidates in N-protein of Viral Haemorrhagic Septicaemia Virus (VHSV): virulence
variability within VHSV Ib clones. Diseases of Aquatic Organisms. 2018; 128:51-62. https://doi.org/10.3354/dao03215

127. Duesund H, Nylund S, Watanabe K, Ottem KF, Nylund A. Characterization of a VHS virus genotype III isolated from rainbow trout (*Oncorhynchus mykiss*) at a marine site on the west coast of Norway. Journal of Virology. 2010; 7(19). doi: 10.1186/1743-422X-7-19

128. Zhu RL, Zhang QY. Determination and analysis of the complete genome sequence of Paralichthys olivaceus rhabdovirus (PORV) Archives of Virology. 2014;159(4):817-820. doi: 10.1007/s00705-013-1716-5.
| Isolate Name       | Genogroup | Host Species | Location                      | Year | Gen Bank       | Citation                      | Coordinates | G-gene haplotype | Combined haplotype |
|-------------------|-----------|--------------|-------------------------------|------|----------------|-------------------------------|-------------|------------------|-------------------|
| DK-Hededam        | I         | North sea cod| North Sea, Denmark            | 1972 | Z93412         | [125]                         | 56.648, 9.271 | N/A              | N/A               |
| French strain 07-71| Ia        | Rainbow trout| Seine-Maritime, France        | 1971 | AJ233396       | Yan et al. unpublished       | 49.9070, 0.7926 | " "              | " "              |
| DK-3592B          | " "       | " "          | North Sea, Denmark            | 1986 | KC778774       | [101]                         | 57.7376, 10.4160 | " "              | " "              |
| De-Fil3           | " "       | " "          | Baltic Sea, Germany           | 1999 | Y18263         | [121]                         | 54.4212, 11.3928 | " "              | " "              |
| Cod-Ucleus         | Ib        | North sea cod| Denmark                       | 1979 | Z93414         | [125]                         | 56.648, 9.271   | " "              | " "              |
| SE-SVA-1033-3F    | " "       | " "          | Kattegatt, Sweden             | 1998 | AB839748       | [126]                         | 56.9793, 12.2100 | " "              | " "              |
| SE-SVA-1033-9C    | " "       | " "          | " "                           | " "  | AB839747       | " "                           | " "          | " "              | " "              |
| SE-SVA-14-3D      | " "       | " "          | " "                           | " "  | AB839745       | " "                           | " "          | " "              | " "              |
| SE-SVA-14-5G      | " "       | " "          | " "                           | " "  | AB839746       | " "                           | " "          | " "              | " "              |
| SE-SVA-1033       | " "       | " "          | " "                           | 2000 | FJ460591       | [111]                         | " "          | " "              | " "              |
| KRRV9601          | " "       | Olive flounder| Seto Inland Sea, Japan        | 1996 | AB672614       | [113]                         | 33.7252, 132.5249 | " "              | " "              |
| DKp37             | " "       | Blue whiting  | North Sea, Denmark            | 1997 | FJ460590       | [111]                         | 57.18766, 8.29058 | " "              | " "              |
| DK-1p49           | II        | Atlantic herring| Baltic Sea                   | 1996 | KM244767       | [118]                         | 55.1418, 15.3020 | " "              | " "              |
| 23-75             | III       | Brown trout   | Eure, France                  | 1975 | FN665788       | [42]                          | 49.2235, 1.2993  | " "              | " "              |
| 14-58             | " "       | Rainbow trout | France                        | 1990 | AF143863       | [110]                         | " "          | " "              | " "              |
| GH40              | " "       | Greenland halibut| Newfoundland                 | 1994 | KM244768       | [118]                         | 47.0017, -44.9999 | " "              | " "              |
| 4p168             | " "       | Atlantic herring| Skagerrack, Denmark           | 1996 | AB672616       | [113]                         | 57.7376, 10.4160 | " "              | " "              |
| FA281107          | " "       | Rainbow trout | Storfjorden, Norway           | 2007 | EU481506       | [127]                         | 62.4056, 6.0446  | " "              | " "              |
| BV860408-52       | " "       | " "          | " "                           | 2008 | FJ362510       | " "                           | " "          | " "              | " "              |
| KRRV9822          | IVa       | Olive flounder| Kagawa, Japan                 | 1998 | AB179621       | Byon et al. unpublished       | 34.3739, 133.9117 | " "              | " "              |
| JF00Ehi1          | " "       | Japanese flounder| Ehime, Japan                  | 2000 | AB490792       | [113]                         | 33.7334, 132.6460 | " "              | " "              |
| FYeosu05          | " "       | " "          | South Korea                   | 2005 | KF477302       | [117]                         | 34.7533, 127.6585 | " "              | " "              |
| Paralichthys       | " "       | " "          | China                         | " "  | KC685626       | [128]                         | 37.5297, 122.1231 | " "              | " "              |
| oliveaceus rhabdovirus (VHSV) | | | | | | | | | |
| KJ2008            | " "       | " "          | Jeju, Korea                   | 2008 | JF792424.1     | [25]                          | 33.4936, 126.5242 | " "              | " "              |
| JF-09             | " "       | " "          | " "                           | 2009 | KM926343       | [102]                         | 33.4936, 126.5242 | " "              | " "              |
| FP-VHS2010-1      | " "       | " "          | Geoje, Korea                  | 2010 | KP334106       | [112]                         | 34.8965, 128.6190 | " "              | " "              |
| ADC-VHS2012-10    | " "       | " "          | Jeju, Korea                   | 2012 | KY979950       | " "                           | 33.4936, 126.5242 | " "              | " "              |
| ADC-VHS2012-11    | " "       | " "          | " "                           | " "  | KY979951       | " "                           | " "          | " "              | " "              |
| ADC-VHS2012-5     | " "       | " "          | Gyeongbuk, Korea              | " "  | KY979946       | " "                           | 36.2894, 128.9379 | " "              | " "              |
| ADC-VHS2012-6     | " "       | " "          | Jeju, Korea                   | " "  | KY979947       | " "                           | 33.4936, 126.5242 | " "              | " "              |
| ADC-VHS2012-7     | " "       | " "          | " "                           | " "  | KY979994       | " "                           | " "          | " "              | " "              |
| ADC-VHS2012-9     | " "       | " "          | " "                           | " "  | KY979949       | " "                           | " "          | " "              | " "              |
| ADC-VHS2013-1     | " "       | " "          | " "                           | " "  | KY979952       | " "                           | " "          | " "              | " "              |
| ADC-VHS2013-2     | " "       | " "          | " "                           | " "  | KY979953       | " "                           | " "          | " "              | " "              |
| ADC-VHS2013-3     | IVa       | Olive flounder| Jeju, Korea                   | 2013 | KY979954       | [80]                          | 33.4936, 126.5242 | N/A              | N/A              |

Table 1. VHSV isolates for full genome analysis. IVb isolates analyzed and their corresponding G-gene and combined gene haplotypes.
| Code       | Taxonomy                | Location                                      | Year   | GenBank Accession | Notes     |
|------------|-------------------------|-----------------------------------------------|--------|-------------------|-----------|
| ADC-VHS2013-4 | Yellow perch            | L. St. Clair, USA                            | 2003   | MK782990          | a         |
| ADC-VHS2013-9 | Yellow perch            | L. St. Clair, USA                            | 2003   | MK783000          | e         |
| ADC-VHS2014-2 | Yellow perch            | L. St. Clair, USA                            | 2004   | MK783000          | e         |
| ADC-VHS2014-4 | Yellow perch            | L. St. Clair, USA                            | 2004   | MK783000          | e         |
| ADC-VHS2014-5 | Yellow perch            | L. St. Clair, USA                            | 2004   | MK783000          | e         |
| ADC-VHS2015-2 | Yellow perch            | L. St. Clair, USA                            | 2005   | MK783000          | e         |
| ADC-VHS2015-5 | Yellow perch            | L. St. Clair, USA                            | 2005   | MK783000          | e         |
| ADC-VHS2016-1 | Yellow perch            | L. St. Clair, USA                            | 2006   | MK783000          | e         |
| ADC-VHS2016-2 | Yellow perch            | L. St. Clair, USA                            | 2006   | MK783000          | e         |
| C03MU*     | Muskellunge             | L. St. Clair, USA                            | 2003   | MK782990          | a         |
| E06FD      | Freshwater drum         | L. Erie, USA                                 | 2006   | MK783014          | This study|
| E06WA      | White bass              | L. Michigan, USA                             | 2006   | MK782986          | N/A       |
| E06WBa     | Yellow perch            | L. Michigan, USA                             | 2006   | MK782986          | N/A       |
| E06YPa     | Smallmouth bass         | L. Ontario, USA                              | 2006   | MK783013          | a4        |
| E06Ypb     | Yellow perch            | L. Michigan, USA                             | 2006   | MK782990          | a         |
| E06Ypc     | White bass              | L. Michigan, USA                             | 2006   | MK782986          | N/A       |
| E06WBb     | Smallmouth bass         | L. Ontario, USA                              | 2006   | MK783013          | a4        |
| C066NP     | Northern pike           | L. St. Clair, USA                            | 2006   | MK782990          | This study|
| C066GS     | Gizzard shad            | L. St. Clair, USA                            | 2006   | MK782990          | This study|
| C066RB     | Rock bass               | L. St. Clair, USA                            | 2006   | MK782990          | This study|
| C066SR     | Shorthead redhorse      | L. St. Clair, USA                            | 2006   | MK782990          | This study|
| C06YP      | Yellow perch            | L. St. Clair, USA                            | 2006   | MK782990          | N/A       |
| C06FD      | Freshwater drum         | L. St. Clair, USA                            | 2006   | MK782990          | N/A       |
| B07RG      | Bluegill                | Budd L., MI, USA                             | 2007   | MK783000          | f         |
| B07PS      | Pumpkinseed             | L. St. Clair, USA                            | 2007   | MK783000          | d         |
| E07CC      | Common carp             | L. Erie, USA                                 | 2007   | MK783000          | g         |
| E07YPa     | Yellow perch            | L. Erie, USA                                 | 2007   | MK783000          | a         |
| E07Ypb     | Yellow perch            | L. Erie, USA                                 | 2007   | MK783000          | a         |
| M075SB     | Smallmouth bass         | L. Michigan, USA                             | 2007   | MK783000          | N/A       |
| M08RB      | Rock bass               | L. Michigan, USA                             | 2008   | MK783010          | i         |
| E08ES      | Emerald shiner          | L. Erie, USA                                 | 2008   | MK783012          | a         |
| E08FDa     | Freshwater drum         | L. Erie, USA                                 | 2008   | MK782993          | a         |
| E08FDb     | Freshwater drum         | L. Michigan, USA                             | 2008   | MK782992          | N/A       |
| M08AM      | Amphipod                | L. Michigan, USA                             | 2008   | MK782990          | This study|
| C08LEa     | Leech                   | L. St. Clair, USA                            | 2008   | MK782990          | This study|
| C08LEb     | Leech                   | L. St. Clair, USA                            | 2008   | MK782990          | This study|
| M08YP      | Yellow perch            | L. Michigan, USA                             | 2009   | MK783007          | j         |
| C09MU      | Muskellunge             | L. St. Clair, USA                            | 2009   | MK782990          | a         |
| Code  | Breed                  | Location                  | Year  | Accession | Genotype | Coordinates       | Location Code |
|-------|------------------------|---------------------------|-------|-----------|-----------|--------------------|---------------|
| M11YP | Yellow perch           | L. Michigan, USA          | 2011  | MK782991  | " "      | 43.0360, -87.8530  | a9            |
| E12FD | Freshwater drum        | L. Erie, USA              | 2012  | MK783004  | " "      | 41.4530, -82.7260  | x             |
| O13GS | Gizzard shad           | L. Ontario, USA           | 2013  | KY359355  | [45]     | 43.2362, -77.5345  | bc            |
| E14GS | Round goby             | L. Erie, USA              | 2014  | KY359356  | " "      | 42.4060, -79.3381  | bd            |
| E15RG | Gizzard shad           | " "                      | 2016  | MK783011  | " "      | 41.7652, -81.2816  | v             |
| E16GSa| Gizzard shad           | " "                      |       | MK782997  | " "      | " "                | " "           |
| E16GSc| Round goby             | " "                      |       | MK782996  | " "      | " "                | " "           |
| E16Gsd| Round goby             | " "                      |       | MK782995  | " "      | " "                | " "           |
| M16RGa| Round goby             | L. Michigan, USA          |       | MK783001  | " "      | 42.9969, -87.8825  | x             |
| M16RGb| Muskellunge            | L. St Clair, USA          | 2003  | MK782981  | " "      | Cell culture       | N/A           |
| CellC03| Gizzard shad           | Lake Erie                 | 2016  | MK782998  | " "      | " "                | " "           |
| Cell16a| Largemouth bass       | " "                      |       | MK783002  | " "      | " "                | " "           |
| Cell16b| Largemouth bass       | " "                      |       | MK782999  | " "      | " "                | " "           |
| Cell16c| Largemouth bass       | " "                      |       | MK782997  | " "      | " "                | " "           |
Table 2. Additional Rhabdovirus sequences used in phylogenetic trees.

| Isolate Name | Host          | Location         | Year | GenBank   | Sequence Citation |
|--------------|---------------|------------------|------|-----------|------------------|
| X89213       | Rainbow trout | Oregon, USA      | 1969 | X89213    | [121]            |
| WRAC         | Chinook salmon| Idaho, USA       | 1994 | L40883    | [119]            |
| " "          | " "           | " "             | " " | NC_001652 | " "             |
| 220-90       | Rainbow trout | Idaho, USA       | 1990 | GQ413939  | [109]            |
| HLJ-09       | " "           | China            | 2009 | JX649101  | [122]            |
| Ch20101008   | Brook trout   | " "             | 2010 | KJ421216  | [114]            |
| BjLL         | Rainbow trout | " "             | 2012 | MF509592  | [124]            |
| NC_000903    | Snakehead murrel | Thailand         | 1988 | NC_000903 | [115]            |
| AF147498     | " "           | " "             | " " | AF147498  | [21]             |
| CA 9703      | Japanese flounder | Japan          | 1984 | NC_005093 | [116]            |
| " "          | " "           | " "             | " " | AF104985  | " "             |
| 80113        | Stone flounder| China            | 2008 | FJ376982  | [123]            |
Table 3. Single nucleotide polymorphisms (SNPs) from each gene’s coding region and the combined non-coding regions (NCDS) in IVb sequence variants of (A) VHSV-IVb, (B) IVa, and (C) Ia, Ib, and III combined. The number of nucleotides (NT) is reported in front of the number of amino acid (AA) changes (the latter are in parentheses). The proportion of nonsynonymous (dN) to synonymous (dS) changes, number of transversions (Tv), and the proportion of transversions to transitions (Ts) are given, along with the evolutionary rate. Totals are in the final row.

**A. IVb**

| Region | Length NT(AA) | # changes NT(AA) | % changes NT(AA) | dN/dS | #Tv | Tv/Ts | Evolutionary Rate |
|--------|---------------|-----------------|-----------------|-------|-----|-------|------------------|
| N-gene | 1215(405)     | 28(10)          | 0.023(0.025)    | 0.357 | 7   | 0.250 | 5.83E-05         |
| P-gene | 669(223)      | 13(3)           | 0.019(0.014)    | 0.231 | 2   | 0.154 | 7.18E-05         |
| M-gene | 606(202)      | 17(9)           | 0.028(0.045)    | 0.529 | 2   | 0.118 | 7.93E-05         |
| G-gene | 1524(508)     | 37(18)          | 0.024(0.035)    | 0.486 | 6   | 0.162 | 8.51E-05         |
| Nv-gene| 369(123)      | 14(7)           | 0.038(0.057)    | 0.500 | 4   | 0.286 | 9.76E-05         |
| L-gene | 5955(1985)    | 112(38)         | 0.019(0.019)    | 0.339 | 18  | 0.161 | 5.02E-05         |
| NCDS   | 745(N/A)      | 32(N/A)         | 0.043(N/A)      | N/A   | 3   | 0.094 | 1.40E-04         |
| Total  | 11083(3446)   | 253(85)         | 0.023(0.025)    | 0.166 | 42  | 0.166 | 6.64E-05         |

**B. IVa**

| Region | Length NT(AA) | # changes NT(AA) | % changes NT(AA) | dN/dS | #Tv | Tv/Ts | Evolutionary Rate |
|--------|---------------|-----------------|-----------------|-------|-----|-------|------------------|
| N-gene | 54(17)        | 1215(405)       | 0.044(0.042)    | 0.315 | 6   | 0.111 | 3.14E-04         |
| P-gene | 45(18)        | 669(223)        | 0.067(0.081)    | 0.400 | 10  | 0.222 | 1.22E-03         |
| M-gene | 36(14)        | 606(202)        | 0.059(0.069)    | 0.389 | 5   | 0.139 | 4.03E-04         |
| G-gene | 87(25)        | 1524(508)       | 0.057(0.049)    | 0.287 | 12  | 0.138 | 5.35E-04         |
| Nv-gene| 31(10)        | 369(123)        | 0.084(0.081)    | 0.323 | 3   | 0.097 | 1.07E-03         |
| L-gene | 424(119)      | 5955(1985)      | 0.071(0.059)    | 0.281 | 65  | 0.153 | 1.02E-03         |
| NCDS   | 110(N/A)      | 720 (N/A)       | 0.153(N/A)      | N/A   | 30  | 0.273 | 4.09E-03         |
| Total  | 787(203)      | 11058(3446)     | 0.071(0.059)    | 0.258 | 131 | 0.166 | 1.06E-03         |

**C. I/III**

| Region | Length NT(AA) | # changes NT(AA) | % changes NT(AA) | dN/dS | #Tv | Tv/Ts | Evolutionary Rate |
|--------|---------------|-----------------|-----------------|-------|-----|-------|------------------|
| N-gene | 1215(405)     | 149(43)         | 0.123(0.106)    | 0.289 | 34  | 0.228 | 4.59E-05         |
| P-gene | 669(223)      | 73(30)          | 0.019(0.135)    | 0.411 | 13  | 0.178 | 4.07E-05         |
| M-gene | 606(202)      | 65(15)          | 0.107(0.074)    | 0.231 | 10  | 0.154 | 5.45E-05         |
| G-gene | 1524(508)     | 198(40)         | 0.130(0.079)    | 0.202 | 41  | 0.207 | 4.54E-05         |
| Nv-gene| 369(123)      | 76(27)          | 0.206(0.220)    | 0.355 | 14  | 0.184 | 7.78E-05         |
| L-gene | 5955(1985)    | 618(70)         | 0.104(0.035)    | 0.113 | 106 | 0.172 | 3.58E-05         |
| NCDS   | 788(N/A)      | 113(N/A)        | 0.143(N/A)      | N/A   | 36  | 0.319 | 3.60E-05         |
| Total  | 11126(3446)   | 1292(225)       | 0.116(0.065)    | 0.174 | 254 | 0.197 | 4.09E-05         |

Table 4. Positive (diversifying) or negative (purifying) selection pressures on individual codons determined by FUBAR (fast, unconstrained Bayesian approximation) and MEME (mixed effects model of evolution) analyses [73, 120] for (A) IVb, (B) IVa, and (C) I/III. None of the codons
found in IVa or I/III match the codons under selection for IVb. Results of seven or more codons 
under selection not displayed. \( pp \)= posterior probability.

**A. IVb**

| Gene | FUBAR diversifying \((pp > 0.95)\) | FUBAR purifying \((pp > 0.95)\) | MEME diversifying \((p<0.05)\) |
|------|----------------------------------|-------------------------------|-------------------------------|
| \( N \) | 0 | 1 (313) | 0 |
| \( P \) | “” | 0 | “” |
| \( M \) | “” | “” | “” |
| \( G \) | 2 (103, 431) | 1 (342) | 1 (431) |
| \( Nv \) | 1 (25) | 0 | 0 |
| \( L \) | 0 | 6 (8, 119, 333, 460, 1284, 1758) | “” |

**B. IVa**

| | | |
| --- | --- | --- |
| \( N \) | 0 | 8 (none in common) | 0 |
| \( P \) | “” | 1 (113) | “” |
| \( M \) | “” | 2 (145, 166) | “” |
| \( G \) | 1 (12) | 5 (28, 75, 157, 216, 301) | “” |
| \( Nv \) | 0 | 1 (24) | “” |
| \( L \) | “” | 30 (none in common) | 3 (147, 593, 1154) |

**C. I+III**

| | | |
| --- | --- | --- |
| \( N \) | 1 (46) | 20 (none in common) | 0 |
| \( P \) | 0 | 8 (none in common) | 0 |
| \( M \) | 0 | 1 (160) | 0 |
| \( G \) | 0 | 19 (none in common) | 1 (477) |
| \( Nv \) | 0 | 3 (56, 96, 109) | 0 |
| \( L \) | 0 | 447 (matches IVb at 1758) | 3 (112, 474, 1012) |
Table 5. Pairwise genetic divergences among VHSV population groupings: (A) sampling time periods, Early (2003–6), Middle (2007–11), and Later (2012–present) and (B) Great Lakes regions (Lake Michigan and Budd Lake, Lake St. Clair, Lake Erie and Lake Ontario) using exact tests (GENEPOP; above diagonal) and $\theta_{ST}$ divergences (ARLEQUIN; below diagonal).

* = $p \leq 0.05$, ** = remained significant ($p < \alpha$) following sequential Bonferroni correction, NS = $p > 0.05$.

| A. Time Groups | B. Regional Groups | L. St. Clair (N=10) | L. Erie and Ontario (N=24) | L. Michigan and Budd (N=9) |
|----------------|--------------------|---------------------|-----------------------------|----------------------------|
| Early (N=16)   | Mid (N=16)         | Late (N=11)         |                             |                            |
| Early          | -                  | NS                  | **                          |                            |
| Mid            | 0.001              | -                   | **                          | **                         |
| Late           | 0.218*             | 0.197*              | -                           |                            |

L. St. Clair

L. Erie and Ontario

L. Michigan and Budd

0.264**

0.348**

-
**Supplementary Table A.** Single nucleotide polymorphisms (SNPs) and nonsynonymous changes per individual isolates. *=Group, includes C06NP, C06RB, C06SR, C06YP, C06FD, M08AM, C08LEa, C08LEb, and C09MU.

| Isolate   | N-gen | P-gen | M-gen | G-gen | N-gen | L-gen | NCDS | Total | Average |
|-----------|-------|-------|-------|-------|-------|-------|------|-------|---------|
|           | NT    | AA    | NT    | AA    | NT    | AA    | NT   | AA    | NT      |
| E06FD     | 1     | 0     | 0     | 0     | 0     | 0     | 3    | 0     | 4       |
| E06WA     | 0     | 0     | 0     | 0     | 1     | 1     | 2    | 1     | 1       |
| E06Wb     | 0     | 0     | 0     | 0     | 0     | 0     | 3    | 1     | 1       |
| E06YPa    | 0     | 0     | 0     | 0     | 1     | 2     | 0    | 0     | 1       |
| E06SB     | 1     | 0     | 0     | 1     | 1     | 1     | 0    | 2     | 0       |
| E06Yb     | 0     | 0     | 0     | 0     | 0     | 0     | 2    | 1     | 0       |
| E06Wbb    | 0     | 0     | 0     | 0     | 1     | 1     | 3    | 3     | 0       |
| O06SB     | 1     | 0     | 0     | 2     | 2     | 6     | 3    | 1     | 11      |
| C06NP*    | 0     | 0     | 0     | 0     | 0     | 0     | 1    | 1     | 0       |
| C06GS     | 1     | 0     | 0     | 0     | 0     | 0     | 3    | 1     | 16      |
| B07BG     | 0     | 0     | 0     | 0     | 0     | 2     | 0    | 1     | 0       |
| B07PS     | 1     | 0     | 0     | 0     | 0     | 0     | 3    | 2     | 1       |
| E07CC     | 1     | 0     | 0     | 1     | 1     | 2     | 1    | 0     | 1       |
| E07Ypa    | 0     | 0     | 0     | 0     | 1     | 2     | 2    | 0     | 2       |
| M07SB     | 1     | 0     | 0     | 0     | 0     | 0     | 3    | 2     | 1       |
| M08RB     | 2     | 2     | 2     | 1     | 0     | 0     | 5    | 3     | 1       |
| E08ES     | 0     | 0     | 0     | 0     | 0     | 0     | 3    | 1     | 1       |
| E08FDa    | 0     | 0     | 0     | 0     | 0     | 0     | 4    | 1     | 1       |
| E08Fdb    | 0     | 0     | 0     | 0     | 0     | 0     | 5    | 2     | 0       |
| M08YP     | 1     | 1     | 0     | 1     | 0     | 0     | 8    | 4     | 12      |
| M11YP     | 0     | 0     | 1     | 1     | 1     | 0     | 3    | 2     | 9       |
| E12FD     | 5     | 2     | 2     | 1     | 2     | 1     | 18   | 6     | 35      |
| O13GS     | 4     | 1     | 2     | 0     | 2     | 1     | 8    | 3     | 9       |
| E14GS     | 2     | 0     | 1     | 0     | 1     | 6     | 2    | 0     | 15      |
| E15RG     | 4     | 2     | 2     | 0     | 1     | 0     | 8    | 4     | 1       |
| E16GsA    | 3     | 1     | 3     | 0     | 1     | 1     | 7    | 7     | 2       |
| E16GsB    | 3     | 1     | 3     | 0     | 1     | 1     | 6    | 6     | 2       |
| E16GSc    | 3     | 2     | 3     | 0     | 1     | 1     | 3    | 3     | 2       |
| E16Gsd    | 3     | 2     | 3     | 0     | 0     | 0     | 3    | 3     | 2       |
| E16Gsc    | 4     | 2     | 3     | 0     | 0     | 0     | 5    | 5     | 2       |
| M16Rga    | 2     | 1     | 1     | 0     | 4     | 1     | 1    | 1     | 1       |
| M16Rgb    | 3     | 1     | 4     | 0     | 3     | 1     | 1    | 1     | 0       |
| Cell16a   | 4     | 1     | 3     | 0     | 1     | 1     | 6    | 6     | 2       |
| Cell16b   | 4     | 1     | 3     | 0     | 1     | 1     | 6    | 6     | 2       |
| Cell16c   | 4     | 1     | 3     | 0     | 1     | 1     | 6    | 6     | 2       |
| CellC03   | 0     | 0     | 0     | 0     | 1     | 1     | 3    | 1     | 0       |

**AVERAGES** 1.553 0.553 1.053 0.105 1.053 0.789 2.816 2.053 0.789 0.395 6.579 2.211 2.289 16.132 6.105 0.064 0.072 0.422
**Figure 1.** Map of VHSV (I–IV) full genome isolates included in this study, colored by Genogroup and with shapes denoting substrains.

**Figure 2.** Map of VHSV-IVb occurrences in the Great Lakes. Shapes colored for time periods (Early: 2003–06, Middle: 2007–10, Late: 2011–17). Diamonds are isolates with sequenced genomes. Large circles depict VHSV-IVb that lack whole genome data.

**Figure 3.** Haplotype network showing genetic relationships among 47 VHSV-IVb full genomes from POPART. Circles sized according to haplotype frequency among isolates. Numbers inside parentheses designate NT differences between each haplotype and the original haplotype, C03MU*. Small, unlabeled black circles = hypothesized haplotype steps.

**Figure 4.** Novirhabdovirus phylogenetic tree, based on full genome sequences (see Tables 3.1 and 3.2), from maximum likelihood and Bayesian analyses. Colored squares designate support values, top left half = Bayesian posterior probabilities, bottom right = 500 bootstrap pseudoreplicates. Tree is rooted to the Snakehead Rhabdovirus (SHRV, GenBank: AF147498).

**Figure 5.** VHSV-IVb phylogenetic tree of IVb whole genome haplotypes, with maximum likelihood and Bayesian analyses. Colored squares = support values, top left half = Bayesian posterior probabilities, bottom right half = 1450 bootstrap pseudoreplicates. Hashes represent cropped region for visualization. *=original IVb isolate. The tree is rooted to substrain IVa (GenBank: JF792424).

**Figure 6.** Phylogenies of individual novirhabdovirus gene, based on full coding sequences of two representatives sequences of all species (see Tables 3.1 and 3.2), with maximum likelihood and Bayesian analyses for (A) NT and (B) AA. Colored squares designate support values, top left half = Bayesian posterior probabilities, bottom right = 2000 bootstrap pseudoreplicates. Tree is rooted to the Snakehead Rhabdovirus (SHRV L-gene, GenBank: AF147498).

**Figure 7.** Host cell viability as measured by Sulforhodamine B (SRB) assay in EPC cells infected with MOIs (MOI=1x10^{-8}–1.0) of four VHSV-IVb isolates, 96 hpi. Average values from
a single experiment (conducted in triplicate) are shown, and are representative of at least three independent experimental replicates. Standard error bars, \( *p<0.05; **p<0.01; ***p<0.001 \).

**Figure 8.** Viral yield assay comparison of infectious viral particles produced (pfu/ml) in wild type (CellC03) and three 2016 VHSV-IVb samples, in BF2 cells 96 hpi, following exposure to media from collection time post infection. Mean values from a single experiment (conducted in triplicate) are shown, and are representative of at least three independent experimental replicates. Standard error bars are shown; \( *p<0.05; **p<0.01; ***p<0.001 \).

**Figure 9.** Antiviral assay comparison of host IFN suppression between reference (CellC03) and three more recent VHSV-IVb isolates, in EPC cells 96 hpi following exposure to UV irradiated media collected at the above time points. Values are quantified as the number of antiviral units (uIFN) per mL. Mean values from a single experiment (conducted in triplicate) are shown, and are representative of at least three independent experimental replicates. Standard error bars, \( *p<0.05; **p<0.01; ***p<0.001 \).

**Figure 10.** Gene expression of host immune response and viral RNA produced. qPCR comparisons between reference (CellC03) and three 2016 VHSV-IVb isolates for (A) EPC IFN and (B) virus detected for samples at each of the above time points. Data were normalized to \( \beta-\)actin mRNA levels. Mean Average values from a single experiment (conducted in triplicate) are shown, and are representative of at least three independent experimental replicates. Significance was calculated from Ct values and transformed values \( (2^{-\Delta\Delta CT}) \) are shown. Standard error bars, \( *p<0.05; **p<0.01; ***p<0.001 \).

**Sup Fig A.** Haplotype network showing genetic relationships among 24 VHSV-IVa full genomes from POPART. Circles sized according to haplotype frequency among isolates. Numbers inside parentheses designate NT differences between each node, unlabeled black circles = hypothesized haplotype steps. Year and location of isolation are below isolate names.

**Sup Fig B.** Haplotype network showing genetic relationships among 16 VHSV-Ia, Ib, and III full genomes from POPART. Circles sized according to haplotype frequency among isolates.
Numbers inside parentheses designate NT differences between each node. Small, unlabeled black circles = hypothesized haplotype steps. Year and location of isolation are below isolate names.
FIGURE 2
FIGURE 4
Figure 5
FIGURE 6A
FIGURE 6B
Cell Viability (% control) vs. Multiplicity of Infection (MOI)

**FIGURE 7**
FIGURE 8

Hours Post Infection (hpi)

Viral Yield (pfu/ml)

- CellC03
- Cell16a
- Cell16b
- Cell16c
FIGURE 10
Sup FIGURE A
