Ohio Sea Grant publication OHSU-TD-139-2013 results of project # R/LR-013

A Dissertation

Entitled

The Evolution and Detection of the Fish Viral Hemorrhagic Septicemia virus (VHSv)

by

Lindsey R. Pierce

Submitted to the Graduate Faculty as partial fulfillment of the requirements for

The Doctor of Philosophy Degree in Biology (Ecology)

_____
Dr. Carol A. Stepien, Committee Chair

_____
Dr. Jonathan Bossenbroek, Committee Member

_____
Dr. Douglas Leaman, Committee Member

_____
Dr. William Sigler, Committee Member

_____
Dr. James Willey, Committee Member

_____
Dr. Patricia R. Komuniecki, Dean
College of Graduate Studies

The University of Toledo

December 2013

An Abstract of

Evolution and Detection of the Fish Viral Hemorrhagic Septicemia virus (VHSv)

by

Lindsey R. Pierce

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Biology (Ecology)

The University of Toledo

December 2013

Understanding how viral populations change over time and space may aid in predicting their future diversification, spread, and success. This dissertation evaluates the evolutionary, biogeographic, and mutational patterns of one of the world's most serious finfish diseases: Viral Hemorrhagic Septicemia virus (VHSv) as a putative emerging quasispecies. Analyses include sequences from five gene regions (glycoprotein – *G*, nucleoprotein – *N*, non-virion – *Nv*, phosphoprotein – *P*, matrix – *M*), and additionally compare detection performance of a newly developed Standardized Reverse Transcriptase Polymerase Chain Reaction (StaRT-PCR) assay. Results suggest that VHSv likely originated from a marine ancestor in the North Atlantic Ocean, diverging an estimated 700 ya into two primary clades: strain IV in North America, and strains I-III in Europe. Strain II appears to comprise the basal group of the latter clade, which differentiated in Baltic Sea estuarine waters. Strains I and III are sister groups, with I found mostly in European freshwaters and III in the marine/estuarine North Sea. Strain IV is differentiated as three monophyletic substrains, with IVa infecting Northeastern Pacific salmonids, IVb endemic to the freshwater Laurentian Great Lakes, and newly-designated IVc in marine/estuarine North Atlantic waters. Overall evolutionary rates of

VHSv appear fastest in the *Nv*-gene ($1.3x10^{-3}$ nucleotide substitutions per site/year), followed by the *G* ($2.6x10^{-4}$), and *N* ($4.3x10^{-4}$). Further examination of IVb identifies 19 unique *G*-gene variants, which evolve quickest in the *Nv*-gene ($2.1x10^{-3}$), followed by the *G* (~1/10 that of *Nv;* $2.1x10^{-4}$), the *P* (~1/15 that of *Nv*; $1.5x10^{-4}$), and the *M*-gene (~1/19 that of *Nv*; $1.2x10^{-4}$). To accurately detect these diversifying variants, a sensitive new StaRT-PCR diagnostic assay was developed, which uniquely incorporates internal standards and can be used on two platforms (Agilent capillary electrophoresis and real-time PCR). Results demonstrate that the StaRT-PCR assays are VHSv-specific to a single molecule (100% accurate at five molecules), control for EDTA and RNA interference, and have no false negative results. By comparison, false negatives ranged from 14-58% in SYBR® green real time RT-PCR tests, and 46-75% in cell culture (the current "gold standard" for VHSv detection). Future research identifying additional sequence variants using StaRT-PCR will aid further understanding of VHSv mutation and spread.

Acknowledgements

This has been a long and worthwhile five-year journey that would not have been possible without several important individuals. First and foremost I would like to thank my advisor Dr. Carol Stepien for the opportunity to work on such an interesting project and for proving support along the way. Thanks are extended to my advisory committee: Drs. Jonathan Bossenbroek, Douglas Leaman, Von Sigler, James Willey, and outside member: Dr. Mohamed Faisal. Your comments and encouragement are appreciated.

Many thanks go to the Great Lakes Genetics/Genomics Laboratory members: Amanda, Susanne, Matt, Vrushalee, Carson, Jhonatan, Tim, and Shane, the George Isaac Cancer Research Center members: Tom, Erin, Lauren, Ji, and Xialou, Dr. Douglas Leaman's laboratory member Adam Pore, and Animal Disease Diagnostic laboratory members Elena Millard and Robert Kim for all of their help. I additionally would like to the Lake Erie Center Staff: Meredith, Rachel, and Pat for their assistance.

It is with immense gratitude that I acknowledge the love and encouragement of my family. You have continued to support my decisions and career path even when they carried me far away from you. Finally, I wish to thank my husband, Russell Friedrich, who has been by my side through both struggles and celebrations. For all of late night talks, for fixing me dinner when I worked 12 hour days, and for all of the encouraging hugs when I didn't feel like I could go on…thank you. Words cannot express my appreciation for all you have done.

Table of Contents

List of Tables

List of Figures

Preface

The chapters of this dissertation are organized into evolutionary analyses and diagnostic development, and are largely identical to their published/submitted manuscript version, with only slight re-wordings.

Chapter 2 previously was published as:

> Pierce, L.R. and Stepien, C.A. (2012) Evolution and biogeography of an emerging quasispecies: Diversity patterns of the fish Viral Hemorrhagic Septicemia virus (VHSv). *Molecular Phylogenetics and Evolution*, **63**, 327-341.

Chapter 3 is planned to be revised and submitted as:

> *Evolutionary Biology*. Stepien, C.A., Pierce, L.R., and Leaman, D.W.

Chapter 4 has previously been published as:

> Pierce, L.R., Willey, J.C., Crawford, E.L., Palsule, V.V., Leaman, D.W., Faisal, M, Kim, R.K., Shepherd, B.S., Stanoszek, L.M., and Stepien, C.A. (2013) A new StaRT-PCR approach to detect and quantify fish Viral Hemorrhagic Septicemia virus (VHSv): Enhanced quality control with internal standards. *Journal of Virological Methods*, **189**, 129-142.

Chapter 5 has been previously published as:

> Pierce, L.R., Willey, J.C., Palsule, V.V., Yeo, J., Shepherd, B.S., Crawford, E.L., and Stepien, C.A. (2013) Accurate detection and quantification of the fish Viral

Hemorrhagic Septicemia virus (VHSv) with a two-color fluorometric real-time PCR assay. *PLoS ONE* 8(8): e71851. doi:10.1371/journal.pone.0071851.

This work was funded by grant awards to CAS from NOAA Ohio Sea Grant Program R/LR-015, USDA-NIFA (CSREES) 2008-38927-19156, 2009-38927-20043, 2010-38927-21048, USDA-ARS CRIS project 3655-31320-002-00D, under the specific cooperative agreement 58-3655-9-748 A01, and NSF-DDIG #1110495 (to LRP, CAS). LRP was supported by NSF GK-12 DGE-0742395 fellowship (P.I.=CAS), and University of Toledo teaching and research assistantships. LRP was awarded the International Association for Great Lakes Research (IAGLR) Scholarship award (2010) and the University of Toledo's Women in STEMM award (2011). This work would not have been possible without the aid of collectors and collaborators from the Ohio and Michigan Departments of Natural Resources, Michigan State University Animal Disease Diagnostic Laboratory (M. Faisal and R. Kim), U.S. Geological Survey in Washington, Seattle (G. Kurath and J. Winton), Cornell University (P. Bowser), University of Wisconsin Milwaukee (F. Goetz), Fisheries and Oceans Canada (K. Garver), the Universidad de Santiago de Compostela (I. Bandín), and the Finnish Food Safety Authority Evira (T. Gadd).

Chapter 1

Introduction

Viruses are some of the most abundant and intriguing infectious agents, which have fascinated scientists and general audiences alike (Holmes, 2009). They have impacted humanity and its food chain for ages, both directly (illnesses) and indirectly (economically by food contamination/loss, etc.). RNA viruses have some of the fastest known evolutionary diversification rates, reaching one nucleotide substitution per round of replication (~6 orders of magnitude faster than their DNA counterparts) (Holmes, 2009). Their small genomes, lack of polymerase proof-reading ability, and short replication times make RNA viruses exemplary model systems to study evolutionary change in "fast forward" (Ojosnegrosand and Beerenwinkel, 2010).

Some RNA viruses diversify into a "cloud" of closely related variants, termed quasispecies (Domingo et al., 2002; Lauring and Andino, 2010). This resultant diversity may allow viral populations to evade host immune responses, improving chances for persistence and spread to new areas and hosts (Jaag and Nagy, 2010). The Viral Hemorrhagic Septicemia virus (VHSv), an RNA rhabdovirus that has impacted fish populations since the 1930s, provides an ideal opportunity to study evolutionary change over time in an important aquatic pathogen.

VHSv is a single-stranded, negative sense RNA rhabdovirus that causes one of the world's most serious finfish diseases, infecting >80 species (OIE, 2013). It is bullet shaped, ~12,000 nucleotides in length, and has six open reading frames of 3'*N-P-M-G-Nv-L*'5. Clinical signs of VHSv infection include internal and external hemorrhaging, bulging eyes, erratic swimming, and fluid filled abdomens (Winton and Einer-Jensen, 2002; Kurath, 2012). Its viral particles (virions) are transported via several natural and anthropogenic vectors, such as fish migration and ballast water transport (Kane-Sutton et al., 2010; Sieracki et al., unpublished data); they thus travel rapidly and widely in aquatic systems.

VHSv first was identified in 1938 in European salmonid aquaculture (described as kidney swelling "*Nierenschwellung* disease"; Schäperclaus, 1938), and since been characterized across the Northern Hemisphere as four genetically and geographically distinct strains (I-IV; Snow et al., 1999). Strain I is classified into five substrains (Ia-e) and predominantly is found in Europe, infecting freshwater salmonids. Strain II is localized in the Baltic Sea, primarily infecting Baltic herring (*Clupea harengus*). Strain III was isolated from several marine fishes in and around the British Isles (Kim and Faisal, 2011). Three substrains of strain IV are endemic to the New World, with IVa along the Northwest Pacific coast (later transported to Asia), strain IVc in the Northwest Atlantic coastal region, and IVb in the freshwater Laurentian Great Lakes. In 2003, IVb emerged in the Great Lakes causing massive fish die-offs (Elsayed et al., 2006). It now is known from all of the Great Lakes, has 17 unique described glycoprotein (*G*) gene variants, and infects >30 species (Groocock et al., 2007; Garver et al., 2013; MEAP-VHSv, http://gis.nacse. org/vhsv/#). The rapid expansion of VHSv into new geographic

areas and host species offers opportunity to evaluate the evolutionary diversification patterns of a putative quasispecies.

Several phylogenetic and genetic distance studies have been conducted since the discovery of VHSv, however its evolutionary and biogeographic relationships remained poorly understood before the present study. The most comprehensive prior description was conducted almost a decade ago by Snow et al. (2004), who included 128 unique *G*-gene sequences from various geographic and host origins. Although that study provided valuable insight into VHSv strains and substrains, it lacked the novel IVb substrain from the Great Lakes. Our inclusion of recent sequence variants was important for understanding how VHSv evolves over time and space, especially in new habitats, such as the Great Lakes.

Increased monitoring efforts to identify VHSv were conducted in response to its 2003 appearance in the Great Lakes. Some fish may be carriers of the disease, but lack visual signs, including external hemorrhages (Bain et al., 2010; Thompson et al., 2011; Cornwell et al., 2012; Faisal et al., 2012).

Disease diagnosis traditionally has relied on the isolation of viral pathogens in cell culture, which is regarded as the "gold standard" (Leland and Ginocchio, 2007). Cell culture is the sole diagnostic for VHSv detection that has been approved by the World Organization for Animal Health (OIE, 2013), along with the joint Fish Health Section of the U.S. Fish and Wildlife Service and the American Fisheries Society (2010). Cell culture relies on confluent cells for viral propagation, which can take up to 28 days to receive results (Garver et al., 2011). It has been shown to lack the sensitivity necessary to diagnose infection in fish harboring low levels of virus (Skall et al., 2005), having false

negatives of 43-95% (Chico et al., 2006; López-Vázquez et al., 2006; Jonstrup et al., 2013). Further, cell culture identification methods, such as plaque assay, may not quantify the number of viral particles.

To circumvent the issues associated with cell culture, molecular diagnostic tools using the polymerase chain reaction (PCR) (Park et al., 2011) have been developed for faster, more sensitive, and accurate disease detection. Ability to diagnose targeted genetic sequences and quantify levels of infectious agents has advanced screening for multiple human diseases, including influenza, hepatitis, and HIV (Coutlée et al., 1991; Ellis and Zambon, 2002). Use of these approaches to elucidate and characterize plant and animal pathogens likewise is growing (Chai et al., 2013; Pasche et al., 2013).

Several PCR-based assays were developed to detect VHSv (Chico et al., 2006; López-Vázquez et al., 2006, Liu et al., 2008, Matejusova et al., 2008, Cutrín et al., 2009; Hope et al., 2010; Garver et al., 2011; Phelps et al., 2012; Jonstrup et al., 2013), however, high false negatives rates (15-92%) were reported (Chico et al., 2006; López-Vázquez et al., 2006; Jonstrup et al., 2013). Those false negatives may have resulted from lack of Internal Standards (IS), which control for interfering substances. IS are recommended by several national (U.S. Environmental Protection Agency, 2004; U.S. Food and Drug Administration 2010) and international agencies (International Organization for Standardization, 2005) and have been implemented in some tests for human diseases, including Hepatitis C virus (Gelderblom et al., 2006) and HIV (Swanson et al., 2006; Schumacher et al., 2007). Thus, a new and improved assay that includes IS was designed here to accurately diagnose and quantify VHSv.

Objectives

The two objectives of this dissertation were to: (1) aid understanding of the evolutionary and biogeographic patterns of VHSv diversification, and (2) develop a rapid and accurate PCR-based assay with IS to detect and quantify the virus. These included four parts, as separate publications:

(1) Evaluate the phylogenetic and biogeographic relationships of the four VHSv strains (I-IV), putative substrains, and isolates using the (*G*), nucleoprotein (*N*), and non-virion (*Nv*) genes. (Chapter 2; Pierce and Stepien, 2012)

(2) Resolve the evolutionary and biogeographic diversification of VHSv substrain IVb in the Great Lakes by analyzing mutational patterns and rates of the *G*, *Nv*, phosphoprotein (*P*), and matrix protein (*M*) genes. (Chapter 3; Stepien and Pierce (with Leaman), under revision, to be submitted)

(3) Develop and evaluate the performance and accuracy of a new VHSv PCR-based assay that incorporates synthetic IS, in comparison to conventional real-time PCR-based tests and cell culture. (Chapter 4; Pierce, Willey, Crawford, Palsule, Leaman, Faisal, Kim, Shepherd, Stanoszek and Stepien, 2013)

(4) Improve the StaRT-PCR assay by making it more user-friendly with real-time PCR equipment. (Chapter 5; Pierce and Willey (with Palsule, Shepherd, Crawford and Stepien), 2013)

Chapter 2

Evolution and biogeography of an emerging quasispecies: Diversity patterns of the fish Viral Hemorrhagic Septicemia virus (VHSv)

Previously published as Pierce, L.R. and Stepien, C.A. (2012) Evolution and biogeography of an emerging quasispecies: Diversity patterns of the fish Viral Hemorrhagic Septicemia virus (VHSv). *Molecular Phylogenetics and Evolution*

2.1 ABSTRACT: Viral hemorrhagic septicemia virus (VHSv) is an RNA rhabdovirus that causes one of the most important finfish diseases, affecting over 70 marine and freshwater species. It was discovered in European cultured fish in 1938 and since has appeared across the Northern Hemisphere. Four strains and several substrains have been hypothesized, whose phylogenetic relationships and evolutionary radiation are evaluated here in light of a quasispecies model, including an in-depth analysis of the novel and especially virulent new substrain (IVb) that first appeared in the North American Laurentian Great Lakes in 2003. We analyze the evolutionary patterns, genetic diversity, and biogeography of VHSv using all available RNA sequences from the glycoprotein (*G*), nucleoprotein (*N*), and non-virion (*Nv*) genes, with Maximum Likelihood and Bayesian approaches. Results indicate that the *G*-gene evolves at an estimated rate of μ=2.58x10^{-4} nucleotide substitutions per site per year, the *N*-gene at μ=4.26x10^{-4}, and *Nv*

fastest at μ=1.25x10^{-3}. Phylogenetic trees from the three genes largely are congruent, distinguishing strains I-IV as reciprocally monophyletic with high bootstrap and posterior probability support. VHSv appears to have originated from a marine ancestor in the North Atlantic Ocean, diverging into two primary clades: strain IV in North America (the Northwestern Atlantic Ocean), and strains I, II, and III in the Northeastern Atlantic region (Europe). Strain II comprises the basal group of the latter clade and diverged in Baltic Sea estuarine waters; strains I and III are sister groups, with the former mostly in European freshwaters and the latter in North Sea marine/estuarine waters. Strain IV is differentiated into three monophyletic substrains, with IVa infecting Northeastern Pacific salmonids and many marine fishes (with 44 unique *G*-gene haplotypes), IVb endemic to the freshwater Great Lakes (11 haplotypes), and a newly-designated IVc in marine/estuarine North Atlantic waters (five haplotypes). Two separate substrains independently appeared in the Northwestern Pacific region (Asia) in 1996, with Ib originating from the west and IVa from the east. Our results depict an evolutionary history of relatively rapid population diversifications in star-like patterns, following a quasispecies model. This study provides a baseline for future tracking of VHSv spread and interpreting its evolutionary diversification pathways.

2.2 INTRODUCTION

RNA viruses, such as the Viral Hemorrhagic Septicemia virus (VHSv), have some of the fastest evolving sequences known – with mutation rates of μ=10^{-4}-10^{-2} nucleotide substitutions per site per year (subst. site^{-1} yr^{-1}) – compared to DNA viruses

that evolve more slowly at $\mu=10^{-8}$-10^{-4} (Li, 1997; Cuevas et al., 2009). RNA viruses evolve quickly due to their small genomes, short generation times, rapid mutation, and lack of polymerase proof-reading (Elena and Sanjuán, 2005). These factors allow some to diversify rapidly from an ancestral sequence into a "cloud" of closely related variants, manifested in a "star-like" pattern on evolutionary trees, which has been termed the quasispecies concept (Domingo et al., 2002; Belshaw et al., 2008; Lauring and Andino, 2010). The resultant diversity of these variants is believed to increase viral population adaptation, survival, and fitness, allowing them to rapidly spread to new hosts and novel environments (Quer et al., 1996; Lauring and Andino, 2010).

The phylogeny and phylogeography of VHSv offers an ideal opportunity to evaluate the patterns of a putative viral quasispecies. Evolutionary and biogeographic relationships among VHSv strains and substrains are poorly known, prior to this investigation. VHSv first was discovered in European aquaculture in 1938, now infects over 70 species (OIE, 2009), and is regarded as one of the world's most important finfish diseases (Wolf, 1988). VHSv is a negative-sense single-stranded RNA rhabdovirus of 11,158 nucleotides, with six open reading frames (3'*N-P-M-G-Nv-L*'5). It has been classified into four strains (I-IV), with several putative substrains, whose relationships are evaluated in this study.

VHSv is most readily transmitted when fish congregate during the spring spawning season in temperatures 9-12^{o}C, with infection leading to erratic swimming behavior, exophthalmia (bulging eyes), bloated abdomens, and extensive external/internal bleeding that results in liver and kidney damage (Winton and Einer-Jensen, 2002). Its viral particles remain infectious for up to 13 days in the water and are transported via

diverse vectors, including: boating, ballast water, fishing tackle, and animals – e.g., amphipod crustaceans, leeches, turtles, and birds (Faisal and Schulz, 2009; Bain et al., 2010; Faisal and Winters, 2011; Goodwin and Merry, 2011). Fish to fish transmission appears prevalent via shed mucus and urine, which likely increases in spawning aggregations (Winton and Einer-Jensen, 2002).

VHSv infection first was discerned in European salmonid aquaculture in the 1930s (Schäperclaus, 1938), and was isolated in 1962 from infected rainbow trout (*Oncorhynchus mykiss*; strain I, isolate DK-F1; Einer-Jensen et al., 2004). It since has been identified across the Northern Hemisphere, with strains I, II, and III predominately occurring in marine/estuarine waters of continental Europe (Fig. 2-1a; Meyers and Winton, 1995; Benmansour et al., 1997; Thompson et al., 2011). Strain I is classified into five substrains: Ia-Ie. Ia is endemic to freshwater European aquaculture, infects 13 fish species, and has the most haplotypes. Ib occurs in marine/estuarine waters of the Baltic and North Seas and infects 10 fish species. Substrain Ie occurs in marine/estuarine Black Sea in brown trout (*Salmo trutta*) and turbot (*Psetta maxima*) (Nishizawa et al., 2006).

Prior to the late 1980s, VHSv was believed to be restricted to continental Europe. In 1988, strain IV was discovered in the North American Pacific Northwest region (now classified as IVa), infecting adult coho (*Oncorhynchus kisutch*) and chinook (*O. tshawytsha*) salmon in hatcheries (Brunson et al., 1989; Hopper, 1989). It then was identified in Pacific cod (*Gadus macrocephalus*) and smelt (*Thaleichthys pacificus*) from Alaska and Oregon in 1999-2001 and Pacific sardine (*Sardinops sagax sagax*) from California in 2002 (Hedrick et al., 2003). IVa is believed to have an extensive reservoir in wild marine fishes in the Northeast Pacific Ocean, including Pacific herring (*Clupea*

harengus pallasi), Pacific hake (*Merluccius productus*), and walleye pollock (*Theragra chalcomgramma*) (Meyers and Winton, 1995; Meyers et al., 1999; Hedrick et al., 2003). Its geographic range then extended to Asia with its isolation from Japanese olive flounder (*Paralichthys olivaceus*) and black seabream (*Spondyliosoma cantharus*) in marine waters of Japan in 1996 and Korea in 1999 (Kim and Faisal, 2010; Studer and Janies, 2011). Strain IV is classified into two substrains: IVa and IVb, along with a third possible group whose relationships are analyzed here.

An especially virulent new substrain – IVb – first was identified from an archived adult muskellunge (*Esox masquinongy*) caught in 2003 from Lake St. Clair of the freshwater North American Laurentian Great Lakes (MI03GL; Elsayed et al., 2006). Outbreaks/occurrences in 2005-2009 infected 31 fish species in all five of the Great Lakes, Lake St. Clair, and inland lakes in Wisconsin, Michigan, and New York (Thompson et al., 2011; see Fig. 2-1b). All but five VHSv-IVb outbreaks to date have been confined to the Great Lakes Basin; the exceptions occurred in Lake Winnebago, WI in 2007, Budd Lake, MI in 2007 and 2009, the Finger Lakes, NY in 2007, Clearfork Reservoir, OH in 2008-2009, and Baseline Lake, MI in 2009 (Bowser et al., 2009; Thompson et al., 2011; Fig. 2-2a,b). The last major Great Lakes' outbreak occurred in 2008 in round goby (*Neogobius melanostomus*) from Lake Michigan near Milwaukee, WI (Wisconsin Aquaculture Assoc., 2011). Since that time, few active outbreaks have occurred; although a variety of Great Lakes fishes have tested positive for VHS, they appeared asymptomatic (Kim and Faisal, 2010). Additionally, VHSv-IVb has been isolated from a leech (*Myzobdella lugubris*) and an amphipod *Diporeia*, which may serve as invertebrate vectors (Faisal and Schulz, 2009; Faisal and Winters, 2011). VHSv-IVb

has been regarded as a significant threat to the world's largest freshwater fisheries in the Great Lakes (Grady, 2007), as well as the baitfish, aquaculture, and tourism industries (Bain et al., 2010; Kim and Faisal, 2010).

Concurrent (2000-2004) VHS outbreaks occurred off the Northwest Atlantic coast of New Brunswick, Canada in marine/estuarine fishes, including mummichog (*Fundulus heteroclitus*), brown trout, striped bass (*Morone saxatilis)*, and three-spined stickleback (*Gasterosteus aculeatus*; Gagné et al., 2007). Partial *G*-gene sequences (724 bp) suggested that these isolates were related to, yet different from IVb (Gagné et al., 2007); their systematic relationships are analyzed here.

Our aim is to evaluate the phylogenetic and biogeographic relationships of the VHS fish virus among all four strains (I-IV), putative substrains, and isolates, comparing all known data for the three genes that have appreciable sequence data at this time: glycoprotein (*G*; 1524 bp), nucleoprotein (*N*; 1214 bp), and non-virion (*Nv*; 366 bp). The *Nv*-gene encodes a small non-virion protein that is known only in fish novirhabdoviruses, e.g., IHNv (Nichol et al., 1995), Hirame rhabdovirus (HIRRv; Kurath et al., 1997), Snakehead (SHRv; Alonso et al., 2004), and VHSv (Thoulouze et al., 2004). We employ Maximum Likelihood (ML) and Bayesian phylogenetic approaches, and evaluate their genetic diversity and comparative divergence times, to provide baseline data for predicting future pathways and understanding their relation to a quasispecies model of evolution. We test the following hypotheses (H_O/H_A): (1) the four strains are not/are each well-supported, reciprocally monophyletic taxa, (2) phylogenies based on the three genes are/are not congruent, (3a) relative rates of evolution are analogous/differ among the four strains, (3b) their evolutionary rates correspond/do not correspond to molecular clock

models, and (4) genetic diversity and levels of divergence are similar/differ within and among the strains.

2.3 MATERIALS AND METHODS

2.3.1 Sequence data and phylogenetic trees

We retrieved all known VHSv sequence data from the literature and NIH GenBank for the *G*-, *N*-, and *Nv*-genes (comprising 613, 205, and 86 sequences, respectively; GenBank, December 20, 2011- http://www.ncbi.nlm.nih.gov/genbank/ and Thompson et al., 2011; see A.2). Locations of isolates were determined from the literature and by our contacting the original authors of the sequences. Unique haplotypes for each gene were identified using ARLEQUIN v3.5.1.2 (Excoffier and Lischer, 2010). To avoid homoplasy, we pruned identical sequences so that each unique haplotype was represented once in our trees; this represented a source of taxonomic confusion in most previous studies. We geographically referenced all sequences and their homologues in A.2. The cDNA haplotypes for each gene were aligned using CLUSTAL X v2.0.12 (Thompson et al., 1997), checked manually, and contiguous sequences were assembled using BIOEDIT v7.0.5 (Hall, 1999).

Phylogenetic trees were analyzed using ML and Bayesian posterior probability (pp) approaches with PHYML v3.0 (Guindon and Gascuel, 2003) and MRBAYES v3.1 (Ronquist and Huelsenbeck, 2003). For the *G*-gene, a partial sequence data set of 669 of 1524 bp was employed to encompass the greatest diversity among available genotype IV isolates, whereas the complete genes for the *N* and *Nv* (1214 and 366 bp, respectively)

were analyzed. An additional analysis of partial *N* sequences (763 bp) was included to encompass newer isolates sequenced from aquatic invertebrates (Faisal and Winters, 2011; see Appendix 2.1). jMODELTEST v0.1.1 (Posada, 2008) evaluated the simplest best-fit model of evolution under a corrected Akaike information criterion (AIC) for each gene. For the *G*-gene, AIC selected the three-phase model with unequal base frequencies (1uf; Kimura, 1981), including invariable sites (i=0.10) and a gamma distribution (α=0.52). Selected best-fit models were: general time reversible (Lanave et al., 1984) plus a gamma distribution (α=0.27) for the *N*-gene, and a three-phase model with unequal base frequencies (2uf; Kimura, 1981) and a gamma distribution (α=0.66) for the *Nv*-gene.

ML nodal support was compared using 2000 nonparametric bootstrap pseudoreplications (Felsenstein, 1985). Bayesian analyses used a Metropolis Coupled Markov Chain Monte Carlo (MC^3) approach, with sampling every 100 of 10 million generations, log likelihood values were plotted at each generation to identify when stationarity was reached, and the first 25% (2,500,000) were discarded as burn-in. Snakehead rhabdovirus (SHRv) was determined by our GenBank Blast searches to be the closest known relative to VHSv (62% similarity; AF147498), and thus was used as the outgroup, as also was done by Ammayappan and Vakharia (2009) and Faisal and Winters (2011).

2.3.2 Total evidence analyses from combined genes

We explored possible combinability of the genes into a single data set using incongruence length difference (ILD; Farris et al., 1994, 1995) in PAUP* v4.0

(Swofford, 2003). Isolates DK-6137, FR-1458, GE-1.2, and UK-8/95 were omitted from this analysis since they lacked *N*- and *Nv*-gene data (~800 bp were missing from *N* and ~100 from *Nv*). Total evidence tests employed 1000 replications in a parsimony framework, with critical α=0.01; since results are susceptible to noise, combinability was analyzed with and without uninformative characters (Cunningham, 1997).

The three gene combined data sets (*G*, *N*, *Nv*) contained 14 VHSv isolates with 3202 aligned bp (*G*-1524 bp, *N*-1214 bp, *Nv*-366 bp), whose GenBank accession numbers are: AB179621, AB490792, AF143862-3, EU481506, FJ460590-91, FN665788, GQ385941, JF792424, NC000855, U28745, U28747, X59241, Z93412, and Z93414. A combined two gene data set (*G* and *N*, which included haplotypes from New Brunswick, Canada) contained 20 isolates of 2736 aligned bp (AB179621, AB490792, AF143862-3, AJ233396, EF079898-9, EU481506, FJ460590-91, FN665788, GQ385941, HQ168405-6, HQ168409, HQ453209-10, JF792424, NC000855, U28747, X59241, Z93412, Z93414). The combined data set omitted strain II since each of its isolates has been sequenced solely for a single gene (excepting DK-1p49, which has been sequenced for the *N*- and *Nv*-genes only). Homologous combined sequences were pruned prior to phylogenetic analysis (these included FJ362510 and FJ362515 for strain III).

Results from the ILD test revealed that the three genes were combinable for a total evidence analysis (p>0.05). jMODELTEST selected the general time reversible model (Lanave et al., 1984) for the combined three and two gene data sets, including gamma distributions (α=0.33 for the former and α=0.29 for the latter). Phylogenetic trees were analyzed for the combined data sets, as described above.

We evaluated patterns of genetic aggregation and spatial dispersion of strain IVb

by constructing a statistical parsimony haplotype network using TCS v1.21 (Clement et al., 2000) for recently published *G*-gene sequence data from Thompson et al. (2011).

2.3.3 Genetic divergences, diversity, and molecular clock calibration

Our PAUP* analyses showed that the individual gene data sets did not conform to a molecular clock model (*G*: χ^2=1042, df=157, p<0.0001; *N*: χ^2=1755, df=23, p<0.0001; *Nv*: χ^2=1397, df=29, p<0.0001). Alternatively, we estimated VHSv divergence times using the Bayesian MCMC approach in BEAST v1.6.2 (Drummond et al., 2002; Drummond et al., 2006), which was developed for analyzing viral strains and has been employed for a variety of viruses (Hughes et al., 2005; Kinnear and Linde, 2010; May et al., 2011). BEAST input gene files were generated using BEAUti v1.5.1 (in BEAST), with the sequence origins here dated to their year of collection/isolation. This method likely resulted in extensive underestimates, with the true sequences origins being older. Our analyses used the ML tree as a starting tree for the MCMC searches, selecting a general time reversible substitution model for the *N*-gene, and the Kimura-2 parameter model for the *G*- and *Nv*-genes, along with an uncorrelated relaxed molecular clock and a coalescent constant size tree prior (Drummond et al., 2006). Searches were run for 10-50 million generations, sampling every 100 generations, with the first 10% (1,000,000-5,000,000) discarded as burn-in. Outputs were assessed in TRACER v1.5 (from BEAST) to ensure that values reached stationarity. The upper and lower 95% highest posterior densities (HPD) for each divergence estimate were reported in TRACER, which provided a Bayesian analog to a confidence interval (Burbrink and Pyron, 2008).

Genetic diversity (h=heterozygosity) and F_{ST} analogues (θ_{ST}; Weir and Cockerham, 1984) of the VHSv strains and substrains were evaluated in ARLEQUIN for the *G*-gene, which contained the most sequence data and diversity. Sequential Bonferroni adjustments were used in all pairwise tests to minimize type I errors (Rice, 1989). For all three genes, number of nucleotide substitutions per site per year (μ=substitutions site^{-1} yr^{-1}), uncorrected pairwise genetic distances (d), and the number of nucleotide differences between (d_{inter}) and among (d_{intra}) strains were calculated in MEGA v5.0 (Tamura et al., 2011). Mismatch distribution analyses, based on the number and distribution of pairwise differences between haplotypes (Rogers and Harpending, 1992) evaluated whether VHSv strains and substrains experienced rapid demographic expansions.

Numbers of transitions/transversions and their relative degrees of saturation were evaluated for 14 VHSv isolates that were fully sequenced for the three genes, representing strains/substrains: I (Z93412), Ia (AF143862-3, FJ460590, FN665788, NC000855), Ib (FJ460591, Z93414), III (EU481506), IVa (AB490792, AB179621, JF792424, U28745, U28747, X59241), and IVb (GQ385941) in MEGA. Other isolates, Y18263 and FJ362510, were available but were homologous to NC000855 and EU481506, and thus were not used. Linear regression (in SPSS v16; SPSS Inc., Chicago, IL; Norusis, 2008) tested fit of the data to a linear model and ANalysis of COVAriance (ANCOVA in R v2.13.1; R Development Core Team, 2012; http://www.r-project.org/) was used to compare models.

Relationships between numbers of non-synonymous (d_N) versus synonymous substitutions (d_S) per nucleotide site were calculated via the Jukes-Cantor (1969) method in SNAP (Synonymous Non-synonymous Analysis Program; http://www.hiv.lanl.gov/

content/sequence/SNAP/SNAP.html; Korber, 2000) for the common isolates sequenced for the three genes. Their ratio – d_N/d_S – (known as the likelihood ratio test) has been used to evaluate potential selection pressures acting on a gene, and to estimate evolutionary divergence (Hall et al., 2005; Hughes et al., 2005). Linear regression models evaluated fit of the data from each gene and ANCOVA was used to compare their relationships. To further test for possible selection, we used Tajima's *D* (Tajima, 1989) and Hudson, Kreitman, and Aguadé (HKA) θ tests (Hudson et al., 1987) in DNAsp v5 (DNA sequence polymorphism; http://www.ub.edu/dnasp/; Librado and Rozas, 2009).

2.4 RESULTS

2.4.1 Genetic variation and divergence of VHSv strains

We evaluated all known VHSv sequences from GenBank and other sources, totaling 904; analyses reveal that 403 are unique, and the rest are synonymized here in A.2. We identify 300 unique *G*-gene haplotypes (of 613 sequences), 67 *N*-gene haplotypes (of 205), and 36 *Nv*-gene haplotypes (of 86). Transitional substitutions significantly outnumber transversions in all three genes: *G* (t=12.8, df=77, p<0.0001), *N* (t=14.2, df=77, p<0.0001), and *Nv* (t=13.0, df=77, p<0.0001), indicating relatively low saturation and consistent phylogenetic signal. The relative number of transitions varies among the *G*- (100), *N*- (64), and *Nv*-genes (39), which significantly differ (F=222.8, df=17, 108, p<0.0001). The relative number of transversions shows a similar trend (*G*-25, *N*-20, *Nv*-11), whose proportions significantly differ among the three genes (F=48.6, df=21, 161, p<0.0001). The relative ratio of transversions:transitions per gene region also

differs (*G*-0.25 Tv:Ts, *N*-0.31, *Nv*-0.28). Although the relationships between the number of transitions to transversions for each gene conform to linear regression models, their relative proportions significantly differ, as shown by their different slopes in Fig. 2-3 and ANCOVA results (F=8130, df=3, 152, $p<2.2\text{x}10^{-16}$).

The 14 common sequences among the three genes do not conform to a molecular clock hypothesis (*G*: χ^2=930, df=11, $p<0.0001$; *N*: χ^2=558, df=11, $p<0.0001$; *Nv*: χ^2=310, df=11, $p<0.0001$). Tajima's *D* tests indicate that the three gene regions approximate neutral models of evolution *(G*: D=0.23, *N*: D=0.13, and *Nv*: D=0.39), congruent with results of HKA tests (*G*: θ=0.07, p=0.99; *N*: θ=0.06, p=0.98; and *Nv*: θ=0.12, p=0.99). All three genes have significantly more synonymous than non-synonymous amino acid changes (F=2818, df=5, 228, $p<2.2\text{x}10^{-16}$), with the *G*-gene having the lowest ratio (d_N=0.02, d_S=0.46, d_N/d_S=0.05), followed by the *N* (d_N=0.03, d_S=0.29, d_N/d_S=0.09), and the *Nv*-gene having the most (mean d_N=0.09, d_S=0.46, d_N/d_S=0.20; Fig. 2-4). More deviations from a linear trend occur in the *Nv*- and *N*-genes (Fig. 2-4), suggesting some positive selection. Evolutionary rates are estimated as $\mu=2.58\text{x}10^{-4}$ for the *G*-gene, $4.26\text{x}10^{-4}$ for *N*-, and $1.25\text{x}10^{-3}$ for the *Nv*-gene. Mismatch distribution analysis supports a rapid range expansion for all four VHSv strains (p=0.67-0.99, NS).

Phylogenetic trees are highly congruent using ML and Bayesian approaches for the *G*-, *N*-, and *Nv*-genes and combined gene analyses (Fig. 2-5). All trees define each of the four strains (I-IV) as reciprocally monophyletic, with high ML (95-100%) and Bayesian (pp=0.90-1.00) support. We observed no evidence for recombination among strains or lateral gene transfer. The primary tree bifurcation (Fig. 2-5a,b,c) distinguishes strain IV from I-III, which diverged an estimated 697 ya (353-1185 ya=95% HPD)

according to the *G*-gene calibration. Within the clade containing strains I-III, II is basal and comprises the sister taxon to the clade of strains I and III; these appear to have diverged ~567 ya (279-941 HPD; Fig. 2-5a). Strains I and III are sister taxa that likely diverged ~403 ya (208-663 HPD).

In the combined gene trees (Fig. 2-5d,e), the primary bifurcation groups strains IV and III as a single clade, which then is the sister group to strain I in the two gene tree (*G* and *N* data; Fig. 2-5e); in the three gene tree, strains I and III group as a single clade that is sister to IV (Fig. 2-5d). Since strain II lacks data for some genes, our analyses could not include it in combined analyses. Thus, our conclusions for strain II are based on the highly congruent and robust results obtained from the separate gene trees, with most divergence estimates based on the *G*-gene, for which more sequence data are available.

Table 2.1 depicts the relative genetic divergences within and among strains. Strains I-IV each have relatively high genetic diversities of h=0.67-0.99 and mean divergence rates of $\mu=1.63\text{x}10^{-4}$ subst. site^{-1} yr^{-1}. Strain I – the first VHSv described strain – is endemic to European marine, estuarine, and freshwaters, has the greatest apparent genetic diversity in our sampling regime (196 *G*-gene haplotypes, h=0.99±0.00, mean inter-strain divergence (d_{inter})=0.14±0.01), and largest intra-strain divergence (mean d_{intra})=0.04±0.004, mean θ_{ST}=0.09, $\mu=1.70\text{x}10^{-4}$). Strain II occurs in the Baltic and Archipelageo Seas and has the second highest genetic diversity (23 *G* haplotypes, h=0.94±0.05, $d_{inter=}$0.14±0.01), but the lowest intra-strain divergence (d_{intra}=0.01±0.00, θ_{ST}=0.12, $\mu=1.28\text{x}10^{-4}$). Strain III appears less diverse, containing 19 *G* haplotypes. It is endemic to the North Sea and has the third highest gene diversity (h=0.85±0.10, d_{inter}=0.14±0.01), with higher intra-strain divergence (d_{intra}=0.02±0.00, θ_{ST}=0.12,

μ=1.42x10^{-4}). Strain IV is found in North America and Asia, has 60 *G* haplotypes, relatively high genetic diversity, and has the most sequence divergence from the other strains, differing by 15 fixed nucleotides (h=0.67±0.03, $d_{inter=}$0.16±0.02). It has the second-most intra-strain divergence (d_{intra}=0.03±0.00, θ_{ST}=0.22, μ=1.78x10^{-4}). Substrains I and IV each have a variety of haplotypes, with wide diversities and divergences; their respective intrastrain patterns are outlined in sections 2.4.2 and 2.4.3, below.

2.4.2 Strain I diversity, relationships, and divergences

Substrains of strain I are estimated to have diverged ~104 ya (61-167 HPD) according to the *N*-gene tree (which is used here, since the *G*-gene phylogeny did not distinguish its substrains clearly; Fig. 2-5a,b). Genetic divergences within and among its five putative substrains (Ia-e) are given in Table 2.1b. Substrains Ia (endemic to Europe) and Ib (Baltic Sea), have higher genetic diversities (Ia: 118 *G* haplotypes, h=0.99±0.00, d_{inter}=0.04±0.01, d_{intra}=0.04±0.00, θ_{ST}=0.06; Ib: 13 haplotypes, h=0.89±0.07, d_{inter}=0.03±0.01, d_{intra}=0.04±0.00, θ_{ST}=0.14); substrains Ic-e lacked sufficient data.

Substrain Ie (marine/estuarine waters of the Black Sea) contains a sole haplotype in our analyses that is located basally on the *G*- and *Nv*-gene phylogenetic trees, appearing as ancestral to the other I substrains. Several strain I isolates, most of which are fairly uncommon, are located next most basally on our trees (Fig. 2-5); these have not been classified into substrains, lack bootstrap support, and show unclear phylogenetic distinctions. The Hededam isolate appears next most basal (to Ie) on the trees, also is marine/estuarine, and infects cultured coho salmon in Denmark. On the *G*-gene tree,

isolate DK-F1 also appears basal and has not been classified as a substrain. The classification of this particular isolate varies between studies, with some grouping it with "I" haplotypes, and others with "Ic" (Einer-Jensen et al., 2004; Duesund et al., 2010). A well-defined cluster of freshwater haplotypes from Finland (FIA, FIP, and FIU; here termed the FIA clade) is supported by our *G*- and *Nv*-gene trees, and appears related to the other strain I unclassified isolates. Substrain Id has been analyzed only for the *G*-gene, contains three haplotypes from Denmark and Finland (only one haplotype shown in our analyses due to lack of sequenced data), and is closely related to the FIA clade (Fig. 2-5a).

Substrain Ib is well supported (97-99% bs/0.97-0.99 pp) in all gene trees (Fig. 2-5 a,b,c). It is marine/estuarine, like most of the other strain I basal groups, and occurs in the North and Baltic Seas. It contains 13 *G*-gene haplotypes and is closely related to the unclassified generic members of strain I (DK-F1, FIA clade) and putative strain Id (Fig. 2-5a). One of the 13 *G*-gene haplotypes was isolated from Japan in 1996 (JP96KRRV9601; in olive flounder). In the *N*- and *Nv*-gene trees, the Cod Ulcus haplotype from the Baltic Sea appears as basal and is quite divergent from the other Ib substrain members (Fig. 2-5b,c). However, it clearly groups with the other Ib isolates in the *G*-gene tree, where it is designated as one of the 13 member haplotypes (Fig. 2-5a).

The *G*-gene sequence tree (Fig. 2-5a) reveals that substrain Ia is paraphyletic, as it groups in three clusters with moderate support (57-79%/0.62-0.82). In our analyses, its subclade of 76 freshwater haplotypes appears monophyletic, with relatively high support (79%/0.82). However, some of these group with other Ia members in the *N*-gene tree (DK3592-B, Fil3, French 0771; Fig. 2-5b) and the *Nv* tree (AU8/95, DK3592-B, Fil3,

DK 6137, French 0771; Fig. 2-5c). In the substrain Ia clade on the *G*-gene tree (Fig. 2-5a), isolates 1458 and FR-2375 are separate from the 76 member clade; however, both 1458 and FR-2375 are found in the main Ia groups on the *N*- and *Nv*-gene trees (Fig. 2-5 b,c). Strain 1458 is located basally in the Ia group in the *G*-, *N*-, and *Nv*-gene trees (Fig. 2-5a,b,c). Other haplotypes within the 76 member clade lacked sufficient data to be included in our combined gene trees. Haplotypes DK-2835, 5123, and 5131 from rainbow trout in Denmark were classified as Ic by Elsayed et al. (2006), but are closely related to Dau 10-97 and Dri 12-95 from Germany, and to AU 62-96 from Austria on our *G*-gene tree (Fig. 2-5a). These six haplotypes lack sufficient data for inclusion in the other gene trees. It may be that these six haplotypes belong to Ia, or that they should be Ic. Additional sequence data are needed to resolve their identity.

2.4.3 Strain IV diversity, relationships, and divergences

In all trees, strain IV forms a clade of three reciprocally monophyletic and supported substrains: IVa, IVb, and our newly designated IVc (100%/0.90-0.99; Fig. 2-5a,b,c). Our trees clearly indicate that IVc merits equal recognition as a valid substrain. The *G*-gene tree depicts IVb as the sister taxon to a clade containing IVa and IVc (62%/1.0; Fig. 2-5a). The *N* tree, for which considerably less sequence data are available, depicts IVa as the sister taxon to a clade containing IVb and IVc (95%/0.95; Fig. 2-5b). IVc sequence data are unavailable, at present, for the *Nv*-gene (Fig. 2-5c). Substrains of strain IV are estimated to have originated ~152 ya (61-283 HPD) (*G*-gene estimate), with substrains a-c appearing to diversify at comparable times (~43 (25-65 HPD), ~29 (11-55

HPD), and ~26 ya (10-50 HPD), respectively).

The three strain IV substrains have high gene diversities, with each showing rapid population expansions. Notably, substrain IVa from the Pacific Northwest and Japan, has the highest comparable genetic diversity (44 haplotypes, h=0.97±0.06, $d_{inter=}$0.04±0.01), and the greatest observed intra-substrain divergence (d_{intra}=0.007±0.00, θ_{ST}=0.18, μ =1.63x10^{-4}; Table 2.1c). We analyzed two *G*-gene isolates from 1996 in Japan (AY167587 and JP99Obama25; both olive flounder), which are located basally within our IVa clade and appear slightly separate from a clade of Northeastern Pacific haplotypes. The two Japanese IVa isolates differ from *G* sequence haplotypes in the Northeast Pacific by 13-21 bp. These distinctions are weakly supported, yet clearly indicate that both strains Ib and IVa separately appeared in the Northwest Pacific region.

The newly-designated substrain IVc (from marine waters of Atlantic coastal Canada) differs from IVb by seven fixed *G*-gene nucleotide substitutions. IVc has the second highest IV substrain genetic diversity in our data set (five haplotypes, h=0.83±0.22, $d_{inter=}$0.03±0.01), and possesses similar intra-substrain diversity to the other IV substrains ($d_{intra=}$0.006±0.001, θ_{ST}=0.22, μ=2.34x10^{-4}).

Substrain IVb is endemic to the freshwater Great Lakes, diverging by 20 fixed nucleotides (11 haplotypes, h=0.67±0.03, $d_{inter=}$0.04±0.01), and its intra-substrain divergence is comparable to the others (d_{intra}=0.004±0.00, θ_{ST}=0.31, μ =1.34x10^{-4}). Relationships among substrain IVb *G*-gene haplotypes are depicted in the haplotype network (Fig. 2-2b). The most common IVb haplotype is MI03GL, which is the original haplotype discovered in 2003 from Lake St. Clair. It has been sequenced 61 times (A.2) and has been found in all five Great Lakes. The second most common IVb haplotype is

U13653 (aka: vcG002), which diverges by a single mutational step from MI03GL; it has been sequenced 36 times and has been isolated from four of the Great Lakes (all except Lake Superior) and from the New York Finger Lakes. Haplotype vcG003 has been found in Lakes Erie and the New York Finger Lakes, whereas the other eight haplotypes each are known from only a single Great Lake. Six haplotypes differ by a single step from MI03GL, and two diverge from it by two steps (vcG008 in Lake Michigan and vcG011 in Lake Erie). The most divergent haplotype is vcG009 from Lake Michigan, which differs by three steps from the geographically widespread U13653 (Fig. 2-2b).

2.5 DISCUSSION

2.5.1 Evolutionary patterns of VHSv: Relation to the quasispecies concept and our hypotheses

Since its 1938 discovery, VHSv displays a diversity of 300 *G*-gene haplotypes, four strains, and multiple substrains, whose star-like radiations follow a quasispecies model showing expansions into new geographic areas, habitats, and hosts. Its rapid diversification and demographic patterns resembles those of other RNA rhabdoviruses. For example, the rabies mammalian RNA rhabdovirus (*Lyssavirus)* that is 35% similar to VHSv (Ammayappan and Vakharia, 2009), was first described in Greece, and then spread across the world (except to Australia and New Zealand). It diverged into two primary clades and seven major subclades (Barbosa et al., 2008), reaching North America in 1958 where it diversified into ~30+ haplotypes within ~53 years (Hughes et al., 2005). Similar diversification characterizes Infectious Hematopoietic Necrosis virus (IHNv;

56% similar to VHSv), which like VHSv, is a *Novirhabovirus* that infects fishes causing hemorrhages (Ammayappan and Vakharia, 2009). IHNv was discovered in the 1950s in the North American Pacific Northwest, and later was detected in Japan, France, and Italy, where it was presumed to have been introduced from North American shipments of infected fish eggs or fry (Kurath et al., 2003). Now, IHNv is classified into four separate strains that span the Northern Hemisphere. It thus shares a general phylogeographic pattern with VHSv substrain IVa.

Our phylogenetic analyses reveal strong support for hypothesis H_A1 that all four VHSv strains comprise reciprocally monophyletic taxa, and for H_O2 that phylogenies based on all three genes support congruent relationships. Unlike our investigation, earlier phylogenetic analyses (e.g., Benmansour et al., 1997; Snow et al., 2004; Einer-Jensen et al., 2005; Elsayed et al., 2006; Nishizawa et al., 2006; Gagné et al., 2007; Lumsden et al., 2007; Ammayappan and Vakharia, 2009; Duesund et al., 2010; Studer and Janies, 2011; Thompson et al., 2011) were based on more limited data sets, did not test gene combinability, almost all did not discern among redundant haplotypes (excepting Einer-Jensen et al., 2004; Snow et al., 2004; Thompson et al., 2011), and all but two did not employ an outgroup to polarize characters (excepting Nishizawa et al., 2002; Faisal and Winters, 2011), and thus were unable to resolve evolutionary direction.

Our evolutionary analyses of the three genes indicate little saturation of phylogenetic signal and support the use of a relaxed molecular clock model (H_A3). The *Nv*-gene appears to evolve the fastest, followed by the *N*-gene (34% of the *Nv* gene's rate), and then the *G*-gene (21% *Nv*'s rate). Genetic diversity levels of VHSv appear relatively high, with strains I and IV having the most haplotypes; comparative genetic

diversity and levels of divergence differ within and among the strains (supporting H_A4). Strain IV appears as the most divergent from the other three strains, with strains I-III being more closely related to each other. Strains I and IV occur in a wide variety of habitat regimes (marine, estuarine, and freshwater), and are more geographically widespread, whereas II and III are exclusively marine/estuarine and appear less diverse (Fig. 2-1a,b). It should be noted that VHSv infections may be relatively cryptic in the wild, especially in marine systems, unless large numbers of infected fishes wash up on shores.

The evolutionary patterns of VHSv fit a quasispecies model in that multiple closely related variants diverge in cluster patterns on our haplotype trees and network, reflecting rapid bursts of evolutionary diversification. These star-like radiations characterize strains I (European marine/estuarine/freshwater), II (Baltic Sea), III (European, marine), and IV (New World and Asia, marine/estuarine/ freshwater). Similar diversification bursts appear to occur within substrains Ia (European freshwaters), Ib (North Sea and Baltic Sea, marine/estuarine), IVa (North American Pacific Northwest and Asia, marine/estuarine /freshwater), IVb (North American Great Lakes, freshwater), and IVc (North Atlantic, marine/estuarine). For example, the substrain IVb network contains 11 haplotypes, with nine diverging in a "cloud" by 1-2 mutational steps from the ancestral haplotype (MI03GL), which was the first isolated sequence and is the most abundant (known from 61 isolates). The variants not only are found in close sequence space, but also occur in close geographic proximity (in the same Great Lake or in neighboring lakes; also see Thompson et al., 2011), supporting a quasispecies model.

2.5.2 VHSv genetic diversity and evolutionary rates

Results of this investigation indicate relatively high genetic diversity and divergence levels of the VHS virus. Transitional substitutions outnumber transversional substitutions in all three genes, with their relative numbers being most disparate in the *G*-gene. Both types of point mutations increase in a relatively linear manner with genetic distance, suggesting the preponderance of low saturation and neutral mutations. Silent (synonymous) substitutions outnumber coding (non-synonymous) amino acid changes in all VHSv genes, with the relative proportion of the latter being appreciably greater in the *Nv*-gene; substitutions in the *Nv*-gene likely are under greater positive selection, meriting additional investigation.

Our evolutionary rate estimate for the *G*-gene is $\mu=2.58\text{x}10^{-4}$, which is slower than the rate estimated by Benmansour et al. (1997) of $1.2\text{x}10^{-3}$ for VHSv, as well as lower than estimates for rabies of $4.10\text{x}10^{-4}$ (Holmes et al., 2002) and $1.2\text{x}10^{-3}$ for IHNv (Kurath et al., 2003). The difference in rate estimates between our study and that of Benmansour et al. (1997) may be attributed to our larger data set (159 sequences of 669 bp) versus the earlier study (11 sequences of only 359 bp). Interpretation of *G*-gene diversity in VHSv-IVb by Thompson et al. (2011) likewise suggests a moderate rate of evolution, in comparison to other *Novirhabdoviruses*.

G-gene divergence rates for VHSv were previously estimated by Einer-Jensen et al. (2004) based on a TipDate model that incorporated a time-structure of serial samples (i.e., isolation date or date of collection), with their substitutions estimated without external calibration (Yang, 2007). As in our study, Einer-Jensen et al. (2004) found that

G-gene sequence evolution did not conform to a molecular clock hypothesis (their study preceded the discovery of substrains IVb and IVc, and later isolates, which are examined here). In response, Einer-Jensen et al. (2004) then divided their analyses into freshwater (Ia) and marine groups (including the remainder of the I substrains, II, III, and IVa), which separately conformed to a molecular clock hypothesis and were calibrated as diverging ~500 ya at μ=1.74x10^{-3} and μ=7.06x10^{-4}, respectively. Our calibrated *G*-gene rate of μ=2.58x10^{-4} thus falls between their two estimates. We attribute these differences in divergence time estimates to the additional sequence information used in our phylogenies and our use of collection/isolation dates, which assumes that VHSv was not present in the environment until the collection date. Additionally, we did not group our data into freshwater and marine strains, as several of the VHSv strains and substrains are euryhaline. Future studies using additional sequence information may offer a more in depth analysis of its evolutionary diversification.

The *G*-gene encodes the primary neutralizing surface antigen for viral attachment and regulates entry into the cell membrane (Plemper, 2011); as the host's immune system continuously evolves its defenses to recognize foreign invaders, the *G*-gene apparently co-evolves in an evolutionary "arms race" to evade the host's immune responses and maintain entry into its cells for infection and replication (Pal et al., 2007). A study of *G*-gene variation in the IV strain found that amino acid variation occurred throughout the length of the protein, but with little variation in the transmembrane regions (Ammayappan and Vakharia, 2009). Einer-Jensen et al. (2004) suggested that nucleotide substitution rates of the *G*-gene in strain I evolved much faster in aquacultured fresh water fish (substrain Ia) than in free-living marine fishes, which may indicate

different selection regimes in different environments. Studies of the related fish *Novirhabdovirus* IHNv identified several sites of likely positive selection in *G*-gene virus surface protein variants (LaPatra et al., 2008). Those changes were hypothesized to possibly result in more efficient binding to the host cells, entry of the virus, or disruption of the interaction of the virus with the proteins of the host immune system. In the latter scenario, specific antibodies or receptors on the cells of host might lower ability to detect the viral variants (Domingo, 2006); in our scenario, this may allow VHSv to evade the fish host's defenses. Responses of host cells to VHSv need further study.

The VHSv *N*-gene appears to evolve at a rate of μ=4.26x10^{-4}, which is similar to the estimated rate for the rabies *N*-gene of 5.27x10^{-4} (Holmes et al., 2002). Thus, the *N*-gene sequences of VHSv and the rabies virus appear to evolve faster than their *G*-genes (~1.7 and ~2.0X faster, respectively). Studies of the rabies virus suggested that the *N*-gene appeared to be largely selective neutral overall when comparing the d_N/d_S ratio (Hughes et al., 2005; Kuzmin et al., 2006), which may be similar to our VHSv *N*-gene findings. The *N*-gene encodes a protein that tightly encloses and protects the RNA in the ribonucleoprotein core from nucleases, and serves as the template for viral transcription and replication (Lou et al., 2007). The *N*-gene's more rapid evolutionary rate may enhance ability of the virus to evade the host's induction of interferon and chemokines, as hypothesized for rabies (Masatani et al., 2010). Our results show that the *N*-gene has an appreciable proportion of non-synonymous substitutions, suggesting that some sites may be undergoing positive selection, which should be investigated further.

The VHSv *Nv* gene appears to evolve faster (μ=1.25x10^{-3}) in comparison to its *G* and *N* VHSv genes (~5X and ~3X faster, respectively). The role of the *Nv* protein may

differ between VHSv and the Novirhabdoviruses IHNv and HIRRv, as hypothesized by Kurath et al. (1997), based on their large sequence divergence. Ammayappan et al. (2010) challenged yellow perch with wild type VHSv-IVb (MI03GL) with the *Nv*-gene intact, versus a modified recombinant strain that lacked the *Nv*-gene; results showed lower rates of fish host mortality, indicating that although the *Nv*-gene is not essential for viral replication, yet appeared to increase replication efficiency and pathogenicity. Investigations of specific amino acid and protein changes of *Nv*-gene variants on host cellular response and relative virulence are needed.

2.5.3 VHSv phylogenetic and biogeographic patterns

Our phylogenetic results clearly distinguish and define all four VHSv strains as well-supported divergent lineages. Our trees show that the primary evolutionary bifurcation separates the ancestors of strains I-III in the Northeastern Atlantic Ocean from the New World group comprising strain IV. This pattern appears consistent with general results of previous studies that were much more limited in scope and lacked character polarization to discern evolutionary direction (e.g., Einer-Jensen et al., 2004, 2005; Snow et al., 2004; Nishizawa et al., 2006; Gagné et al., 2007; Lumsden et al., 2007; Ammayappan and Vakharia, 2009; Duesund et al., 2010; Studer and Janies, 2011; Thompson et al., 2011). The two primary clades in our trees largely are geographically separated, indicating that this is a vicariant pattern. The most basal members of both clades are marine, which supports the origin of VHSv in marine waters, as suggested by other studies (e.g., Dixon, 1999; Snow et al., 2004; Einer-Jensen et al., 2004; Thompson

et al., 2011). Evidence for outbreaks of marine VHSv often may be unobserved, as infected fishes may not wash up onto shores and would be rapidly consumed by predators.

Our phylogenies indicate that the I-III clade then diverged into two groups: strain II, which is found in marine/estuarine waters of the Baltic Sea, and a second clade that unites strains I and III to a common ancestor. Strain II comprises a tight genetic cluster that is confined to a few geographically proximate sites in the Baltic Sea and Archipelago Finland (with only 57 documented occurrences). Our trees suggest that strain III then separated in the North Atlantic Ocean and the North Sea (marine/ estuarine), and largely is peripheral to and encircles the distribution of strain I. Strain III has nine unique *G*-gene haplotypes on our tree. Strain I likely then diverged from an ancestor shared with III, with I radiating widely via European freshwater aquaculture in the 20th century.

Strain IV is the most divergent clade from I-III, and appears to have originated in marine fishes in the Northwest Atlantic Ocean. Our results show that strain IV is diversified in two primary clades: one comprises IVb in freshwaters of the Laurentian Great Lakes, and the other links IVa from the Northern Pacific Ocean and IVc from the Northwest Atlantic Ocean – the latter's identification as a novel substrain is supported here. It is likely that IVb diversified from an unknown marine reservoir in the Northwestern Atlantic Ocean, perhaps near the mouth of the St. Lawrence River. It then likely mutated to diverge with rapid radiation in the Great Lakes, following a quasispecies model, as seen in our data. Strain IV relationships are further discussed below.

2.5.4 Strain I diversification and phylogeographic patterns

Substrains within strain I overlap spatially in Europe and appear to have radiated quickly. Traditionally, strain I has been classified into five substrains (Ia-e; Thiéry et al., 2002), with several marine/estuarine haplotypes remaining unclassified (here designated as "I"). These unclassified isolates are located basally on our trees, having low bootstrap support and little phylogenetic definition; their location suggests that this group has a marine ancestry in the environs of the Northwest Atlantic Ocean. The marine ancestry hypothesis for I also was outlined by Snow et al. (2004) using *N*- and *G*-gene sequences.

Our phylogeny indicates that substrain Ia is well supported, which is divided into three groups in the *G*-gene tree. One of its clades contains 76 closely related *G* haplotypes in another star-like radiation; some of these are located in the monophyletic Ia clades of the *N*- and *Nv*-gene trees (AU 8/95, DK3592-B, DK6137, Fil3, French0771, X73873; others have not been sequenced for the latter genes). These haplotypes include isolates from brown and rainbow trout in European freshwaters of France, Denmark, Germany, Austria, and Slovenia, and one from brackish waters in Finland. The Ia cluster, which includes European isolates Datt107, FR2375, and strain 1458, are found in freshwaters of France and Germany. It may be that the two Ia groups comprise separate substrains; however, many isolates in the 76 haplotype clade either have not been sequenced for the *N*- and *Nv*-genes, or have incomplete data. A third Ia cluster contains freshwater isolates from Germany and Austria (AU62-96, Dau10-97, Dri12-95), which groups closely with Ic (freshwater rainbow trout from Denmark). These Ic haplotypes may merit synonymy with substrain Ia (here labeled as "Ic"/Ia), but lack *N* or *Nv*

sequence data to resolve their classification. An early phylogenetic study by Einer-Jensen et al. (2004) discerned especially high rates of evolution for their Ia isolates from aquaculture.

Substrain Ib is a well-supported separate taxon in our phylogenies. The substrain is believed to have originated from asymptomatic marine fish species and has relatively low pathogenicity (Dixon et al., 1997; Snow et al., 2004). One of its haplotypes – Cod Ulcus (from the Baltic Sea) – appears divergent in the *N* and *Nv* gene trees, meriting further investigation. The Cod Ulcus isolate is contained within the 13 haplotype clade on our *G*-gene tree. Substrain Id is closely related to a group of five haplotypes labeled as "I"/FiA. It may be that this "I"/FiA haplotype group should be grouped into substrain Id, or that the Id haplotype belongs to the FiA taxon. Recent work by Duesund et al. (2010) and Gadd et al. (2011) grouped the FiA clade (including some FiU and FiP isolates) as Id on their *G*-gene trees, however, classification of this group cannot be resolved here, as further *N*- and *Nv*-gene sequence data are needed. On our *G*-gene tree, the Ie isolate GE-1.2 appears closely related to the Hededam isolate from Demark. Nishizawa et al. (2006) examined additional partial *G*-gene sequences (only 359 bp) that were closely related to GE-1.2, which are not used here to due their data limitations. Their data analysis suggested that Ie likely arose from an indigenous VHSv strain in the Black Sea. Additional sequence information for the *G*-, *N*-, and *Nv*-genes are necessary in order to verify this hypothesis.

2.5.5 VHSv in the Northwest Pacific

Our *G*-gene tree indicates that two separate VHS substrains infected Japanese olive flounder in 1996: Ib and IVa. They likely followed different routes to Asia, with Ib traveling eastward from Europe and IVa transported west from the Northeastern Pacific Ocean. These strains (and the Asian isolates) appear at very different locations in our phylogeny. Studer and Janies (2011) also conducted a biogeographic analysis using a subset of the *G*-gene sequences we used and independently came to the conclusion supporting two separate invasions of VHSv to Asia; however, their methods were unable to resolve the phylogenetic relationships. The Asian isolates have not yet been sequenced for the other genes, which may reveal key biogeographic and diversity information. It also is unknown whether the Northwest Pacific Ocean harbors an extensive VHSv reservoir in its marine fishes.

2.5.6 Strain IV diversification and phylogeographic patterns

Our study defines IVc isolates from Northwestern Atlantic Ocean marine and estuarine waters as a valid substrain. Our phylogenies are unclear as to whether IVc is the sister group of IVa (supported by the *G*-gene), or IVb (indicated by the *N* gene); however, both clearly define its substrain validity. Further resolution of these evolutionary patterns should incorporate *Nv* sequence data.

Putative common ancestry of substrains IVa and IVc may trace to the Northwest Atlantic Ocean. It appears possible that a wide-ranging marine fish, invertebrate, or

shipping vector led to the transport of ancestral VHSv-IV across the Arctic from the Northwest Atlantic into the Northeast Pacific Ocean. This is the more likely scenario since IVa has an extensive marine fish reservoir, infecting wild Pacific herring and sardines (Meyers and Winton, 1995; Hedrick et al., 2003). IVa also reportedly has lower virulence than the VHSv strains from Europe and North America that infect freshwater fishes (see Winton and Einer-Jensen, 2002; Hedrick et al., 2003; Studer and Janies, 2011). Alternatively (and less likely), VHSv-IV could have been accidentally transported from the east coast of North America in attempts to establish the Atlantic salmon (*Salmo salar*) in the Pacific Northwest in the 1930s (Washington State Dept. of Health, 2011) or perhaps via transport of various trout species for aquaculture in the early to mid 1900s (Oregon's Agricultural Progress, 2011). For example, aquaculture facilities off the coast of Washington in Pacific marine waters have reared Atlantic salmon in net pens since the early 1950s, which are located in close geographic proximity to where the first IVa isolate was discovered in 1988 (Waknitz et al., 2002). Thus, fish transports might have served as vectors for VHSv-IV.

Since outbreaking in the Great Lakes in 2005, substrain IVb has mutated and diversified in a star-like phylogenetic pattern. Higher diversity of *G*-gene variants is apparent in the Lakes Ontario and Erie regions (see Fig. 2-2b), suggesting multiple introduction and spread pathways. IVb has infected 31 species of fishes to date (Thompson et al., 2011), and has been identified in the leech *Myzobdella lugubris* (Faisal and Schulz, 2009) and an amphipod *Diporeia* (Faisal and Winters, 2011); *G*-gene sequences of the latter were identical to the most common IVb isolate MI03GL. Faisal and Winters (2011) additionally sequenced 5 pooled *Diporeia* samples, identifying two

new *N*-gene haplotypes that differed by 1 and 2 bp from MI03GL (Fig. 2-5b inset, isolates HU54-M and ON41). It will be interesting to sequence additional invertebrate samples in light of possible reservoirs for fish VHSv.

Interestingly, Thompson et al. (2011) found two cases where two different isolates had co-infected single fish samples *in situ,* with each containing the common MI03GL sequence, along with another variant that dominated that particular outbreak. They hypothesized this might be due to both sequences infecting the same fish, or from possible new mutations originating in the host individual (Thompson et al., 2011). Other cases of co-infection with different *G*-gene variants were documented in European VHSv fish hosts (Einer-Jensen et al., 2004), and for the related fish rhabdovirus IHNv in Russia (Rudakova et al., 2007) and Europe (Enzmann et al., 2010). Such diversity in an individual fish host and among members infected in a given geographic outbreak appears to fit a quasispecies model, in which diverse variants are genetically linked through mutation and may interact cooperatively on a functional level, collectively contributing to the success of the viral population (see Lauring and Andino, 2010). Diverse variants may allow VHSv to readily adapt to a variety of new environments and hosts.

Over recent years, increasing numbers of a variety of fishes have tested positive for VHSv-IVb, but have been asymptomatic, indicating that fish hosts have developed resistance (Kim and Faisal, 2010). Over time, there have been fewer reported outbreaks in the Great Lakes, suggesting that IVb has become endemic and less virulent; these hypotheses remain to be investigated. It will be interesting to compare VHSv-IVb diversity patterns in the more rapidly evolving *Nv*-gene.

2.5.7 Summary, conclusions, and recommendations for future work

Interpreting the phylogenetic patterns of VHSv is important to understanding its emergence pathways, diversification, and potential to adapt and spread as an emerging quasispecies. Our study provides a robust phylogenetic analysis, revealing congruent patterns among its *G*-, *N*-, and *Nv*-genes. We clarify the taxonomy of strain I and IV substrains, and provide evidence for IVc as a separate taxon. Future study of the *Nv*-gene's evolution will provide a valuable comparison to the *G*- and *N*-genes, as we show that it evolves more rapidly, possibly functioning in viral replication and pathogenicity (Ammayappan and Vakharia, 2009). The present results indicate that the evolutionary history of VHSv is intriguing, and likely will provide important co-evolutionary insights through its future progression.

2.6 ACKNOWLEDGEMENTS

This study was funded by NSF DDIG #1110495 (for LRP to work with CAS), along with grants to CAS (PI) and colleagues from USDA-NIFA (CSREES) #2008-38927-19156, #2009-38927-20043, #2010-38927-21048, USDA-ARS #58-3655-9-748 A01, NOAA Ohio Sea Grant #R/LR-015, and NSF GK-12 #DGE-0742395 (LRP in 2010-11). We are grateful for the help of D. Leaman, A. Pore, and present and past members of the Great Lakes Genetics Laboratory, including: H. Dean, A. Haponski, S. Karsiotis, D. Murphy, M. Neilson, V. Palsule, O.J. Sepulveda-Villet, and T. Sullivan. We thank P. Uzmann and M. Gray for logistic support, J. Bossenbroek for statistical assistance, Gael Kurath, USGS, Seattle who provided a preprint copy of Thompson et al.

(2011), and J. Tomelleri for permission to use copyright images for graphical abstract. This is publication #2012-003 from the Lake Erie Research Center.

Table 2.1

Pairwise comparisons of genetic variation in the *G*-gene (a) within and among VHSv strains, (b) strain I substrains, and (c) strain IV substrains. Along the diagonal (**bold**) are uncorrected *p*-distances±standard error within strains/substrains; above diagonal are comparisons among the strains/substrains. Below the diagonal are θ_{ST} divergences (Weir and Cockerham, 1984). *=Significant θ_{ST} difference at 0.05 pre-sequential Bonferroni correction (Rice, 1989). **=Remained sig. after correction. N/A=not available, due to lack of sequence data.

(a)

	I	**II**	**III**	**IV**
Strain I	**0.036±0.004**	0.131±0.013	0.112±0.012	0.166±0.015
Strain II	0.029**	**0.006±0.002**	0.126±0.014	0.175±0.017
Strain III	0.070**	0.102**	**0.020±0.003**	0.180±0.018
Strain IV	0.162**	0.223**	0.264**	**0.027±0.004**

(b)

	I	**Ia**	**Ib**	**Ic**	**Id**	**Ie**
Strain I	**0.018±0.003**	0.037±0.005	0.019±0.004	0.027±0.005	0.015±0.003	0.032±0.006
Substrain Ia	0.215**	**0.035±0.004**	0.040±0.006	0.045±0.006	0.038±0.006	0.053±0.007
Substrain Ib	0.338**	0.057**	**0.004±0.001**	0.028±0.006	0.014±0.004	0.031±0.006
Substrain Ic	0.380	0.008	0.075	**0.009±0.003**	0.027±0.006	0.039±0.007
Substrain Id	0.501	0.011	0.115	0.000	**N/A**	0.028±0.006
Substrain Ie	0.501	0.011	0.115	0.000	1.000	**N/A**

(c)

	IVa	**IVb**	**IVc**
Substrain IVa	**0.007±0.002**	0.045±0.008	0.036±0.007
Substrain IVb	0.280**	**0.004±0.001**	0.025±0.006
Substrain IVc	0.085*	0.349**	**0.006±0.002**

Fig. 2-1 Maps showing the distributions of VHSv strains (I-IV, colored), substrains (symbols, Ia-e and IVa-c), and sequenced isolates from: (a) Europe and Asia, and (b) North America. The first isolate and date for each strain are geographically referenced.

Fig. 2-2 Distribution (a) and *G*-gene haplotype network (b) of VHSv-IVb variants in the Laurentian Great Lakes and inland waterbodies (designated by watershed) in TCS v1.21 (Clement et al., 2000). Raw data are from Thompson et al. (2011). Squares in haplotype network are sized according to the number of observed isolates and are colored according to the map. Lines denote a single mutational step between haplotypes; small, unlabeled circles represent hypothesized unsampled haplotypes. Parentheses below the isolate name contain the number of documented occurrences and outbreak year.

Fig. 2-3 Number of transitions (open symbols) and transversions (closed symbols) versus pairwise genetic distances (calculated in MEGA v5) for common isolates sequenced for the three genes: *G* transitions (R^2=0.99, F=34101, df=1, 76, p<0.001) and transversions (R^2=0.99, F=5616, df=1, 76, p<0.001; a), *N* transitions (R^2=0.97, F=2190, df=1, 76, p<0.001) and transversions (R^2=0.97, F=2817, df=1, 76, p<0.001; b), and *Nv* transitions (R^2=0.99, F=10044, df=1, 76, p<0.001) and transversions (R^2=0.97, F=2408, df=1, 76, p<0.001; c). The slopes of all lines significantly vary (F=8130, df=3, 152, $p<2.2\text{x}10^{-16}$). Lack of overlap of transitions and transversions indicates low saturation.

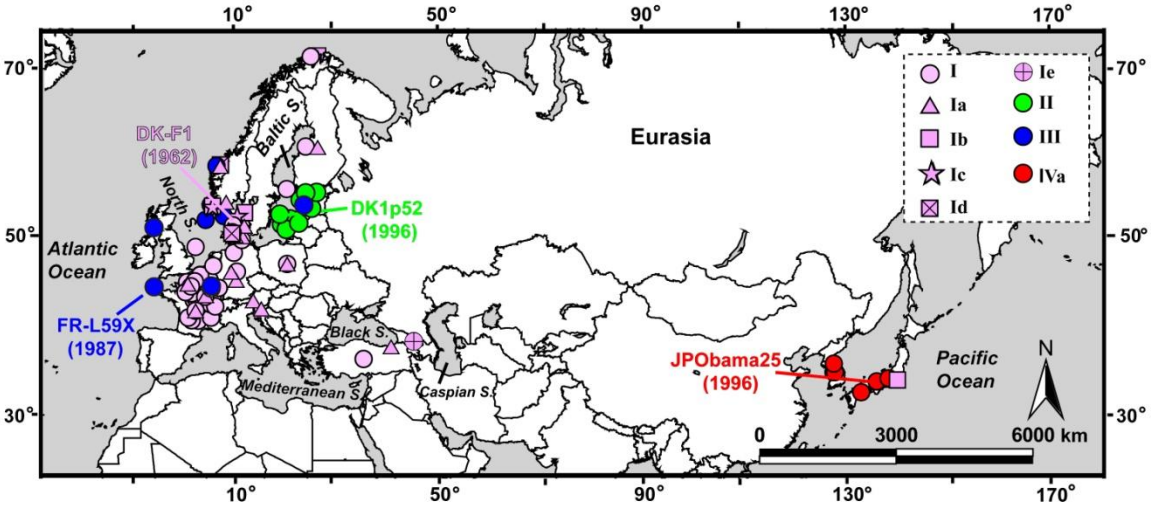
Fig. 2-4 Regression of mean numbers of non-synonymous (d_N) versus synonymous substitutions (d_S) per nucleotide site (calculated with the Jukes-Cantor (1969) method in SNAP) for the common isolates sequenced for the three genes. (Regression equations: *G* – R^2=0.88, F=531, df=1, 76, p<0.001; *N* – R^2=0.92, F=914, df=1, 76, p<0.001; *Nv* – R^2=0.98, F=3646, df=1, 76, p<0.001). The ratio of d_N/d_S significantly varies among the three genes (F=2818, df= 5, 228, $p<2.2\text{x}10^{-16}$).

Fig. 2-5 Maximum Likelihood phylogenetic consensus trees of VHSv sequences for: (a) *G*-gene, (b) *N*-gene – with inset showing substrain IV relationships from partial sequences, (c) *Nv*-gene, (d) all three genes combined, and (e) genes *G*- and *N*- combined.

Trees are congruent with those from our 50% majority rule Bayesian analyses. Numbers in black triangles=number of unique sequences per subgroup. Values above nodes=% support from 1000 bootstrap pseudo-replications/Bayesian posterior probability. Values in parentheses and italics=estimated divergence time (years). Snakehead rhabdovirus (AF147498) was used as the outgroup.

Fig. 2-1

(a)


10°
50°
90°
130°
170°
70°
50°
30°
Eurasia
Atlantic Ocean
Pacific Ocean
Baltic S.
North S.
Black S.
Mediterranean S.
Caspian S.
DK-F1 (1962)
DK1p52 (1996)
FR-L59X (1987)
JPObama25 (1996)
I
Ia
Ib
Ic
Id
Ie
II
III
IVa
N
0
3000
6000 km

(b)

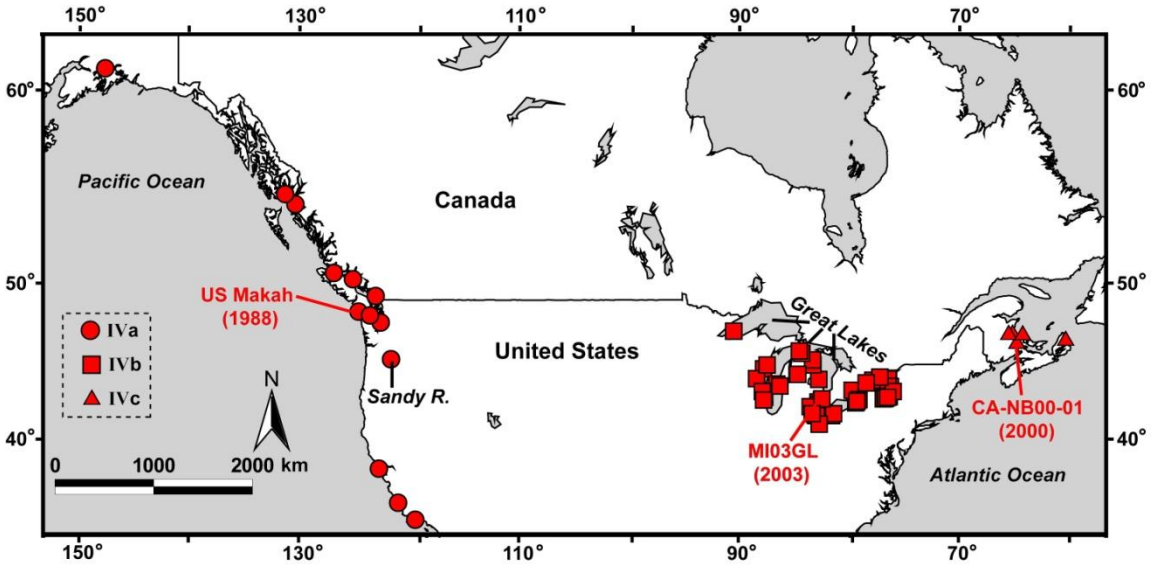
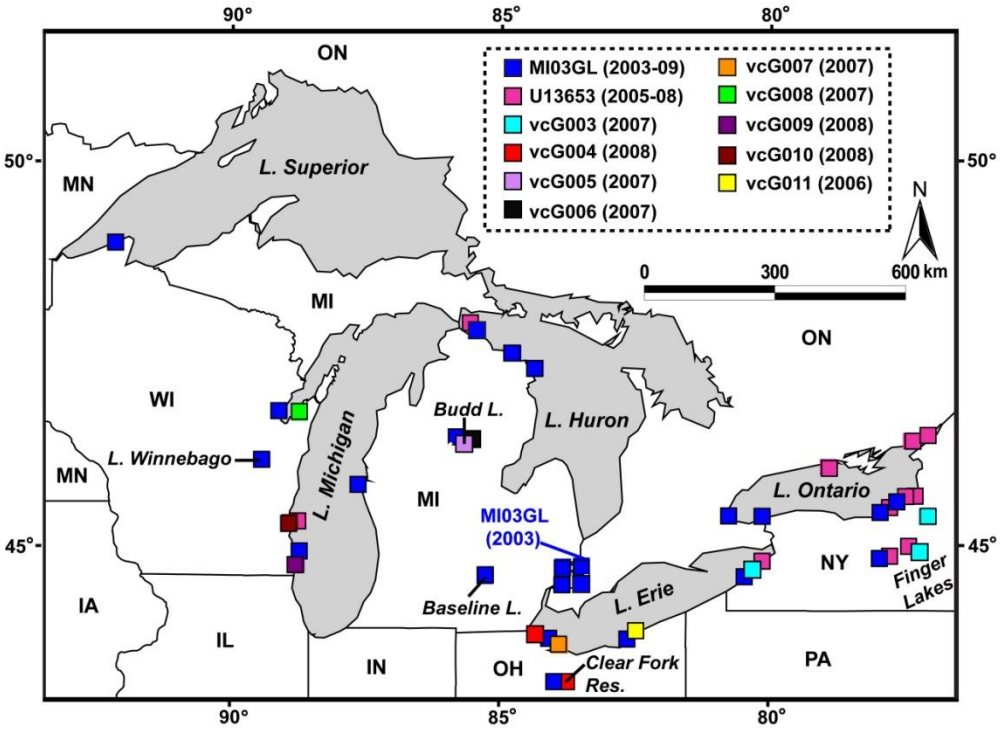

150°
130°
110°
90°
70°
60°
50°
40°
Pacific Ocean
Canada
United States
Atlantic Ocean
Great Lakes
US Makah (1988)
Sandy R.
MI03GL (2003)
CA-NB00-01 (2000)
IVa
IVb
IVc
N
0
1000
2000 km

Fig. 2-2

(a)


90°
85°
80°
50°
45°
ON
MN
WI
MI
IA
IL
IN
OH
NY
PA
L. Superior
L. Michigan
L. Huron
L. Erie
L. Ontario
Budd L.
L. Winnebago
Baseline L.
Clear Fork Res.
Finger Lakes
MI03GL (2003)
MI03GL (2003-09)
U13653 (2005-08)
vcG003 (2007)
vcG004 (2008)
vcG005 (2007)
vcG006 (2007)
vcG007 (2007)
vcG008 (2007)
vcG009 (2008)
vcG010 (2008)
vcG011 (2006)
N
0
300
600 km

(b)

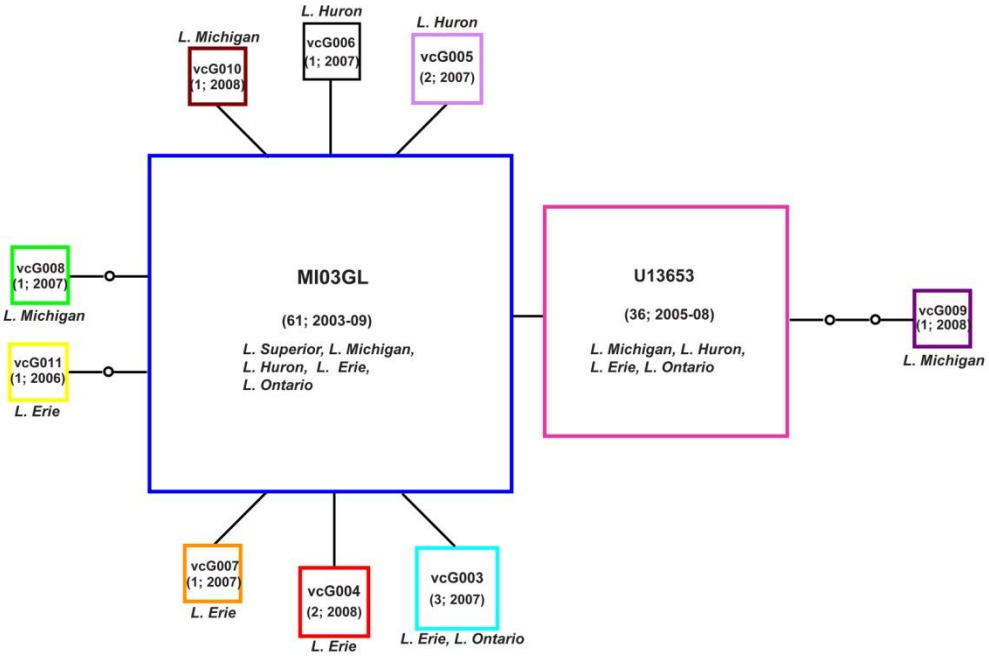

L. Michigan
vcG010 (1; 2008)
L. Huron
vcG006 (1; 2007)
L. Huron
vcG005 (2; 2007)
vcG008 (1; 2007)
L. Michigan
vcG011 (1; 2006)
L. Erie
MI03GL
(61; 2003-09)
L. Superior, L. Michigan, L. Huron, L. Erie, L. Ontario
U13653
(36; 2005-08)
L. Michigan, L. Huron, L. Erie, L. Ontario
vcG009 (1; 2008)
L. Michigan
vcG007 (1; 2007)
L. Erie
vcG004 (2; 2008)
L. Erie
vcG003 (3; 2007)
L. Erie, L. Ontario

Fig. 2-3

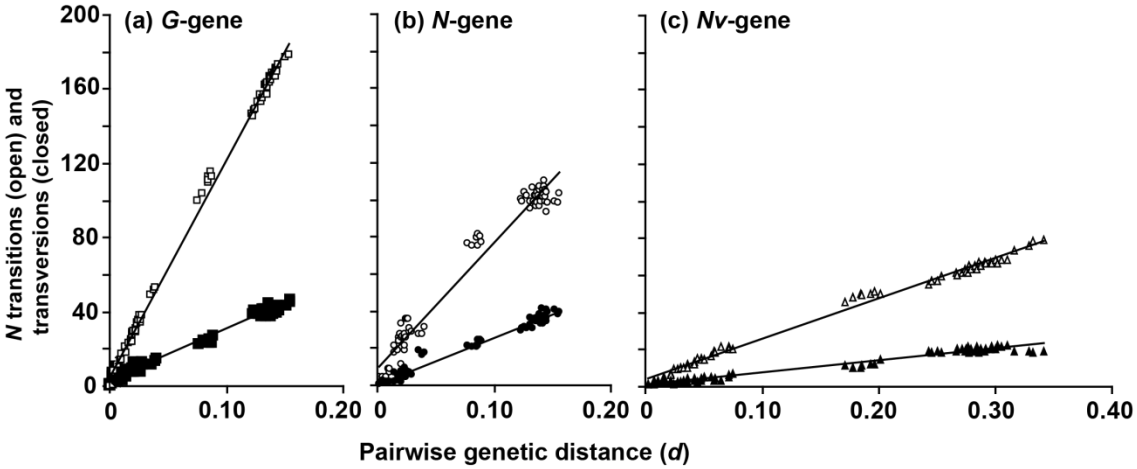

(a) G-gene
(b) N-gene
(c) Nv-gene
N transitions (open) and transversions (closed)
200
160
120
80
40
0
0
0.10
0.20
0.30
0.40
Pairwise genetic distance (d)

Fig. 2-4

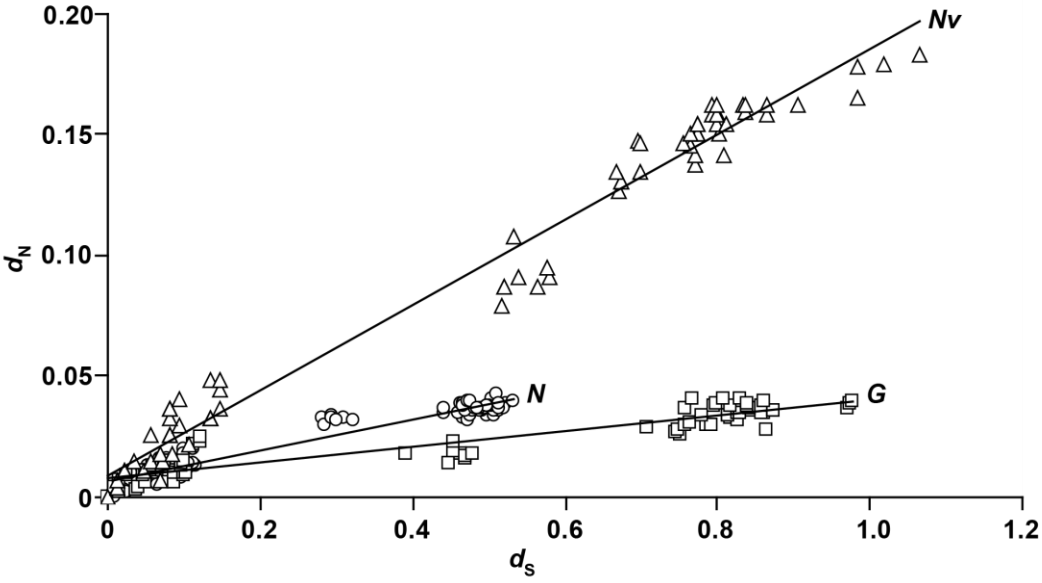


Fig. 2-5 (a)

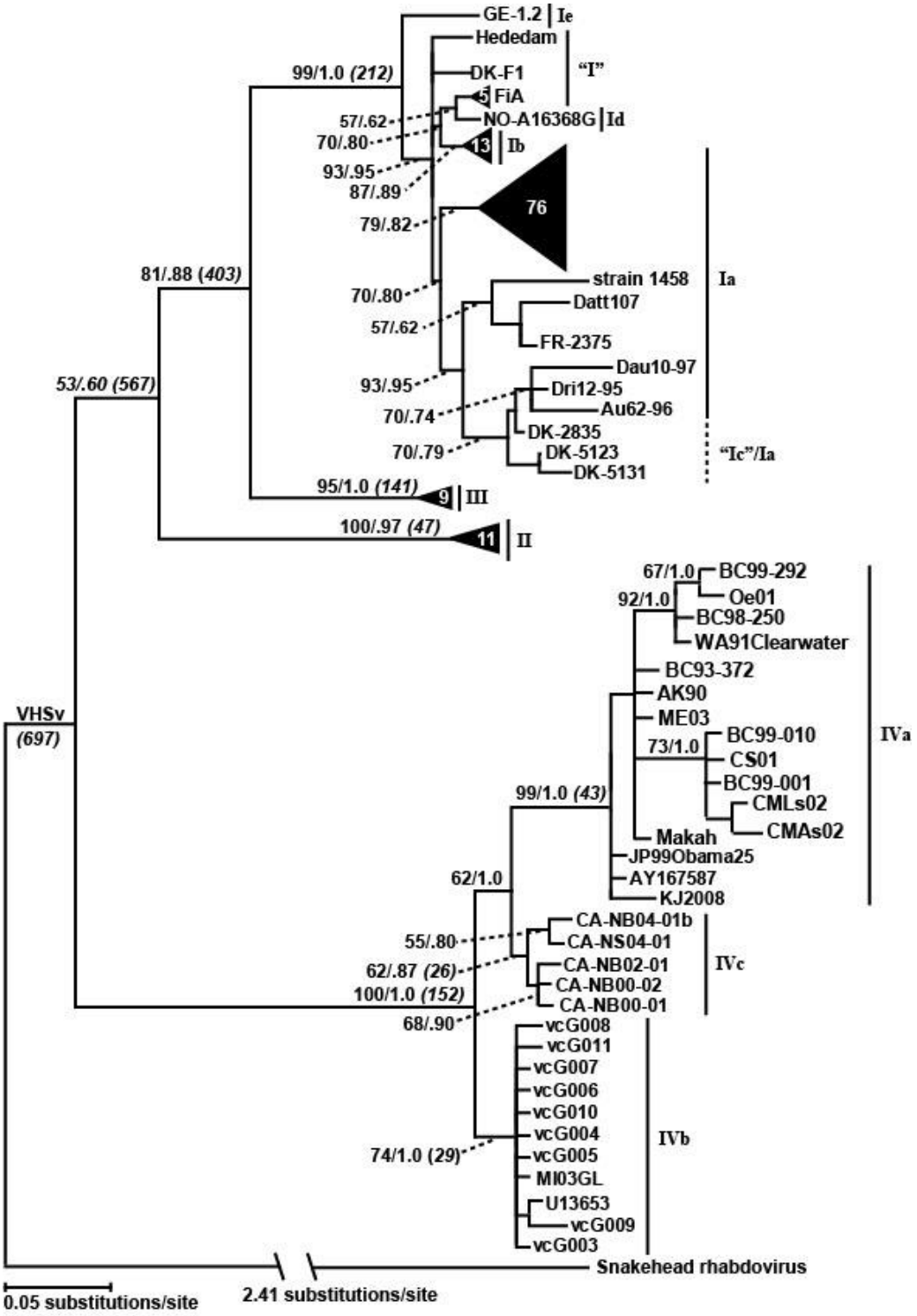

GE-1.2
Ie
Hededam
"I"
99/1.0 (212)
DK-F1
5
FiA
57/.62
NO-A16368G
Id
70/.80
13
Ib
93/.95
87/.89
76
79/.82
81/.88 (403)
strain 1458
Ia
70/.80
Datt107
57/.62
FR-2375
Dau10-97
53/.60 (567)
93/.95
Dri12-95
Au62-96
70/.74
DK-2835
70/.79
DK-5123
"Ic"/Ia
DK-5131
95/1.0 (141)
9
III
100/.97 (47)
11
II
67/1.0
BC99-292
92/1.0
Oe01
BC98-250
WA91Clearwater
BC93-372
AK90
VHSv
(697)
ME03
73/1.0
BC99-010
IVa
CS01
BC99-001
99/1.0 (43)
CMLs02
CMAs02
Makah
JP99Obama25
62/1.0
AY167587
KJ2008
CA-NB04-01b
55/.80
CA-NS04-01
62/.87 (26)
CA-NB02-01
IVc
100/1.0 (152)
CA-NB00-02
CA-NB00-01
68/.90
vcG008
vcG011
vcG007
vcG006
vcG010
vcG004
IVb
74/1.0 (29)
vcG005
MI03GL
U13653
vcG009
vcG003
Snakehead rhabdovirus
0.05 substitutions/site
2.41 substitutions/site

Fig. 2-5 (b)

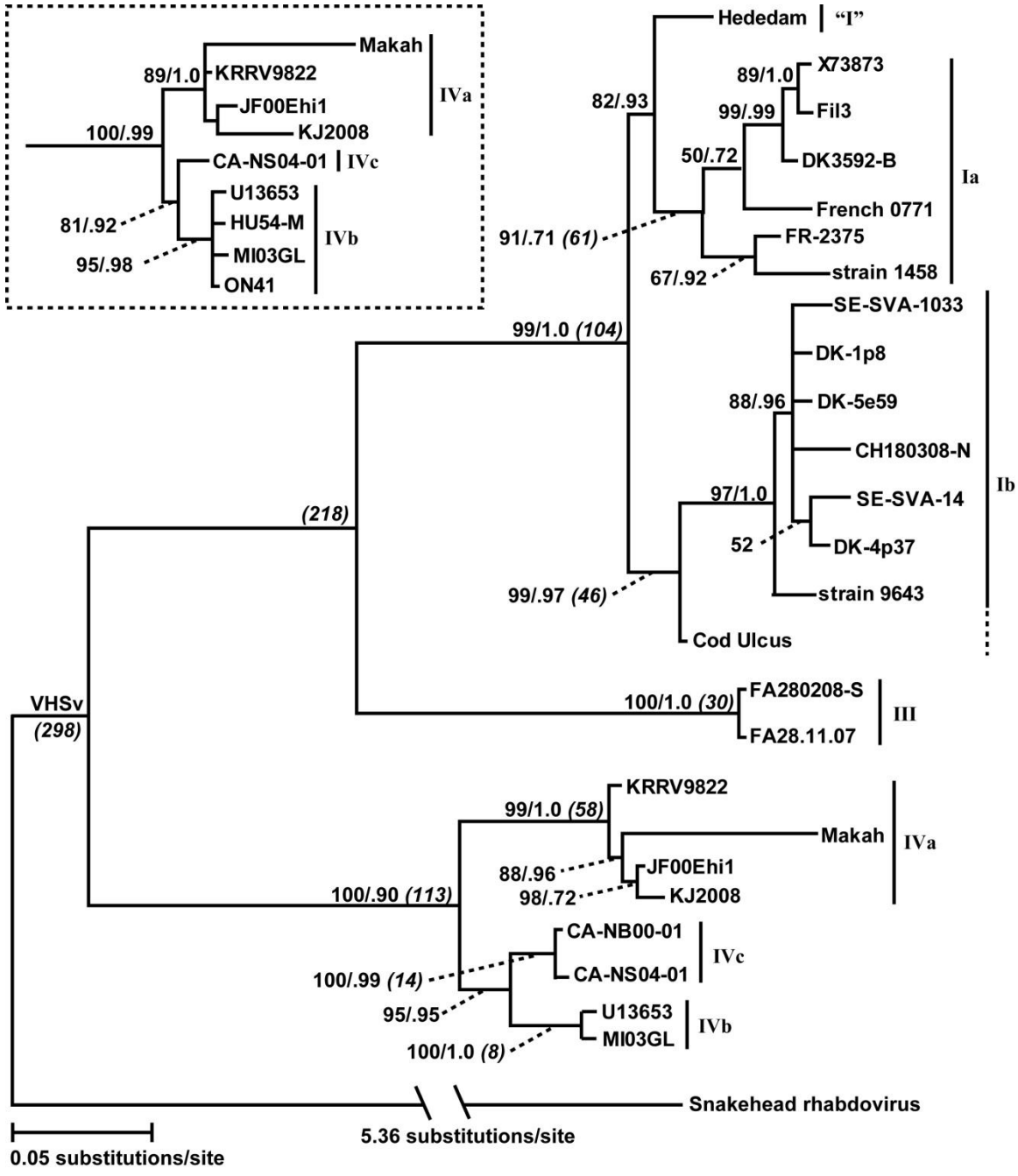

Makah
89/1.0
KRRV9822
JF00Ehi1
KJ2008
IVa
100/.99
CA-NS04-01
IVc
U13653
81/.92
HU54-M
IVb
95/.98
MI03GL
ON41
Hededam
"I"
89/1.0
X73873
82/.93
99/.99
Fil3
50/.72
DK3592-B
Ia
French 0771
91/.71 (61)
FR-2375
67/.92
strain 1458
SE-SVA-1033
99/1.0 (104)
DK-1p8
88/.96
DK-5e59
CH180308-N
Ib
97/1.0
SE-SVA-14
(218)
52
DK-4p37
99/.97 (46)
strain 9643
Cod Ulcus
VHSv
(298)
100/1.0 (30)
FA280208-S
III
FA28.11.07
KRRV9822
99/1.0 (58)
Makah
IVa
88/.96
JF00Ehi1
100/.90 (113)
98/.72
KJ2008
CA-NB00-01
IVc
100/.99 (14)
CA-NS04-01
95/.95
U13653
IVb
MI03GL
100/1.0 (8)
Snakehead rhabdovirus
0.05 substitutions/site
5.36 substitutions/site

Fig. 2-5 (c)

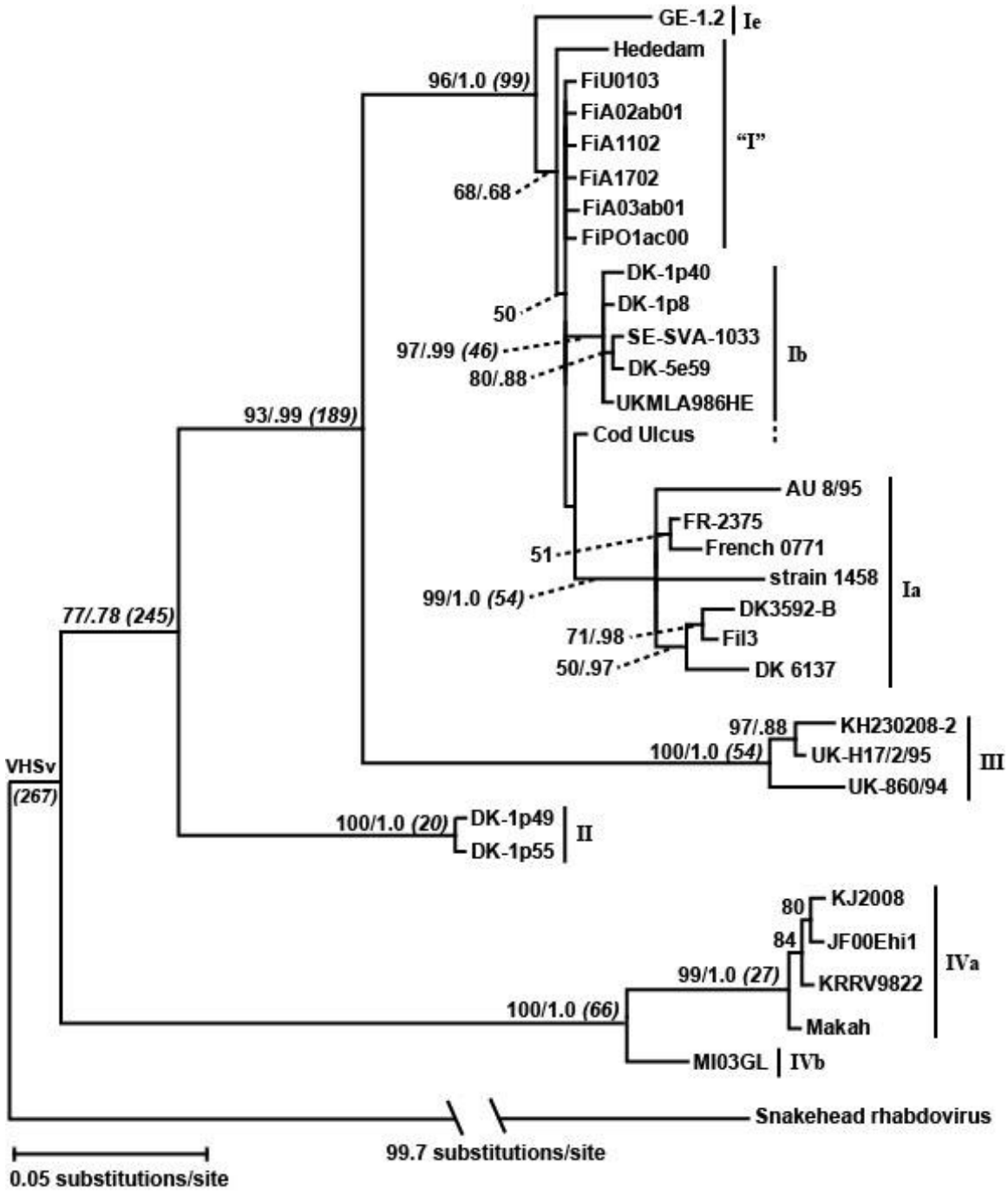

GE-1.2
Ie
Hededam
96/1.0 (99)
FiU0103
FiA02ab01
FiA1102
"I"
FiA1702
68/.68
FiA03ab01
FiPO1ac00
DK-1p40
50
DK-1p8
97/.99 (46)
SE-SVA-1033
Ib
80/.88
DK-5e59
UKMLA986HE
93/.99 (189)
Cod Ulcus
AU 8/95
FR-2375
51
French 0771
strain 1458
Ia
99/1.0 (54)
77/.78 (245)
DK3592-B
71/.98
Fil3
50/.97
DK 6137
97/.88
KH230208-2
100/1.0 (54)
UK-H17/2/95
III
VHSv
(267)
UK-860/94
100/1.0 (20)
DK-1p49
II
DK-1p55
80
KJ2008
84
JF00Ehi1
IVa
99/1.0 (27)
KRRV9822
100/1.0 (66)
Makah
MI03GL
IVb
Snakehead rhabdovirus
99.7 substitutions/site
0.05 substitutions/site

Fig. 2-5 (d)

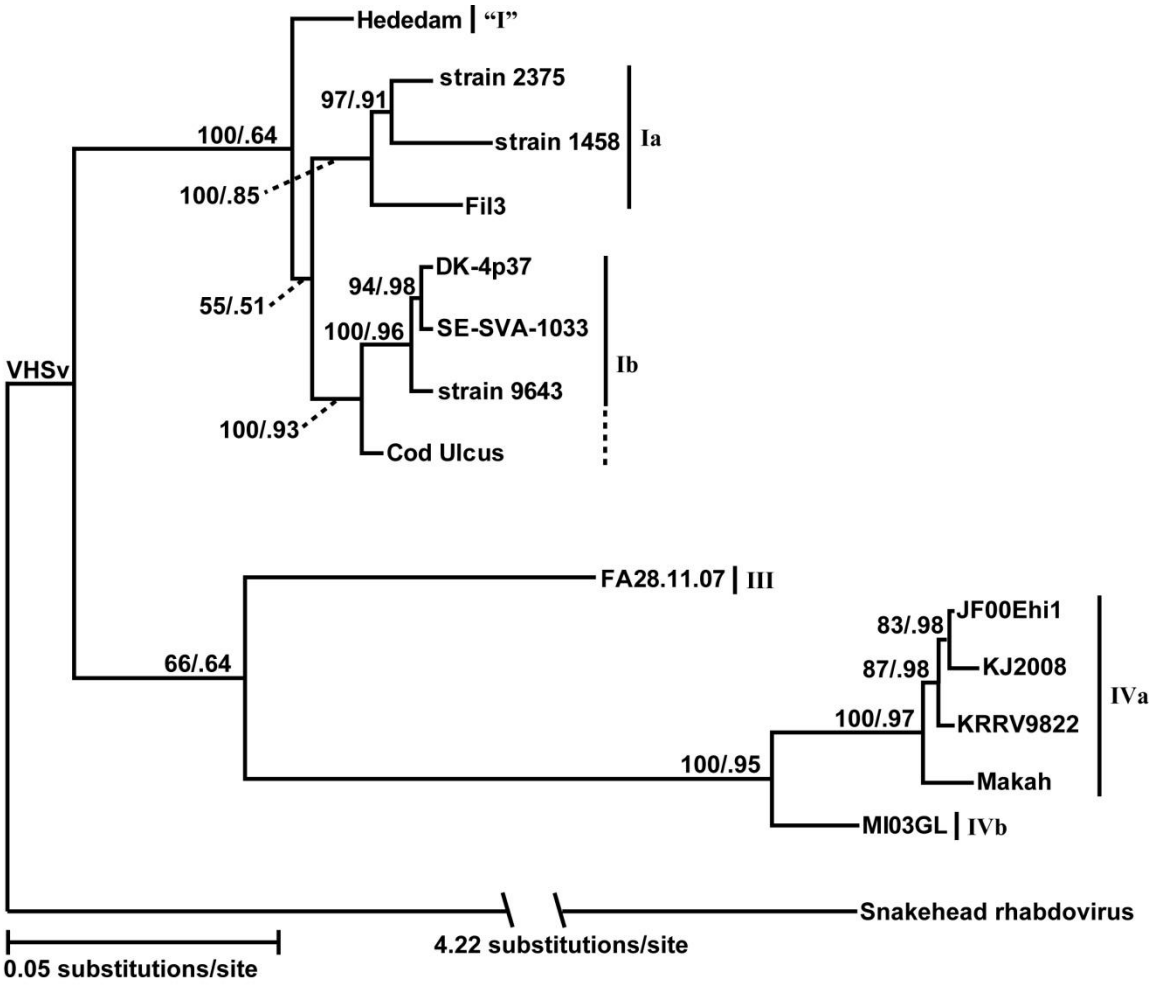
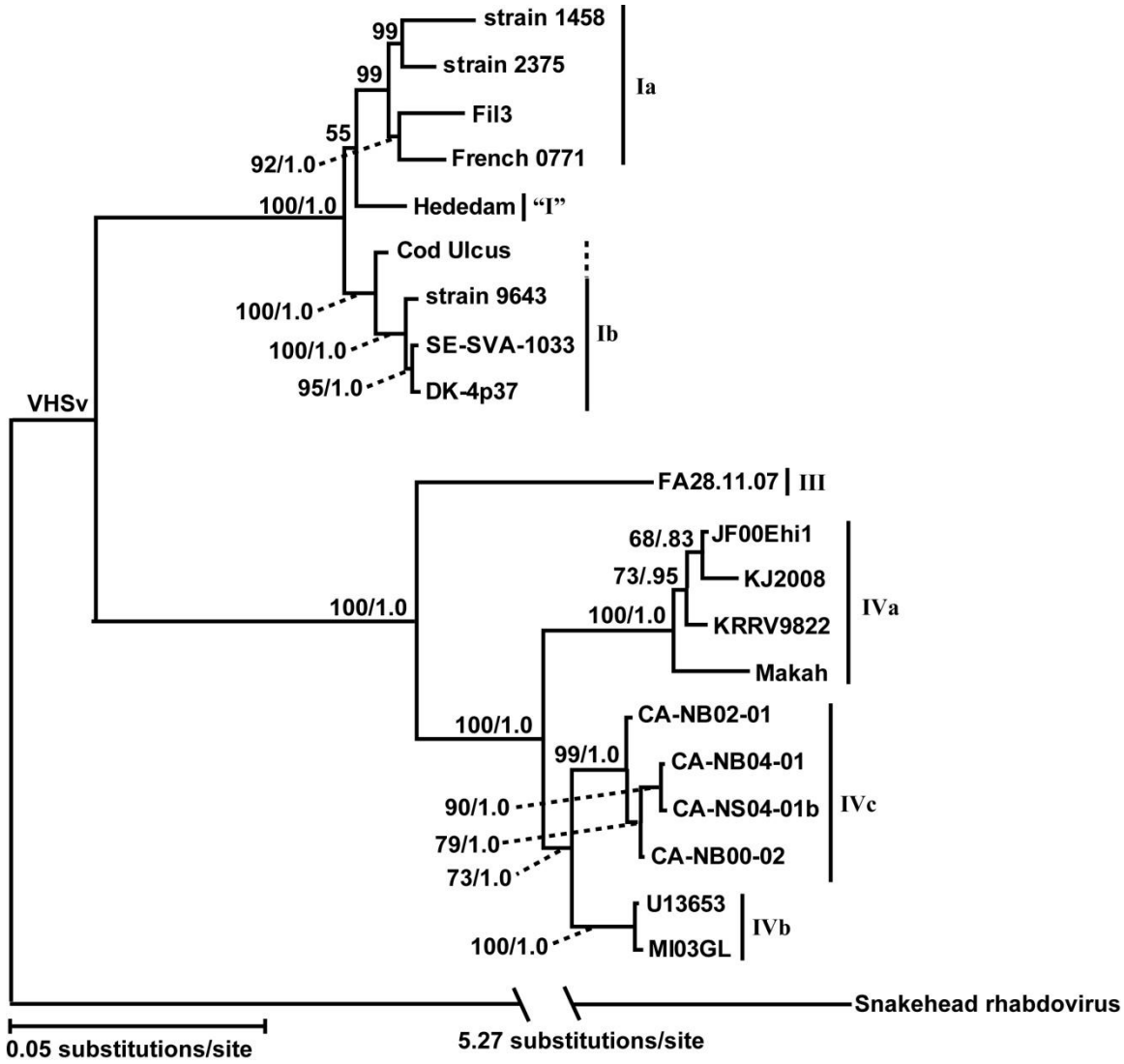

Hededam | "I"
strain 2375
97/.91
strain 1458
Ia
100/.64
100/.85
Fil3
DK-4p37
94/.98
55/.51
SE-SVA-1033
100/.96
Ib
VHSv
strain 9643
100/.93
Cod Ulcus
FA28.11.07 | III
JF00Ehi1
83/.98
KJ2008
87/.98
IVa
66/.64
100/.97
KRRV9822
100/.95
Makah
MI03GL | IVb
Snakehead rhabdovirus
4.22 substitutions/site
0.05 substitutions/site

Fig. 2-5 (e)


strain 1458
strain 2375
Fil3
French 0771
Hededam
"I"
Ia
Cod Ulcus
strain 9643
SE-SVA-1033
DK-4p37
Ib
99
99
55
92/1.0
100/1.0
100/1.0
100/1.0
95/1.0
VHSv
FA28.11.07
III
68/.83
73/.95
100/1.0
JF00Ehi1
KJ2008
KRRV9822
Makah
IVa
100/1.0
100/1.0
CA-NB02-01
CA-NB04-01
CA-NS04-01b
CA-NB00-02
IVc
99/1.0
90/1.0
79/1.0
73/1.0
U13653
MI03GL
IVb
100/1.0
Snakehead rhabdovirus
0.05 substitutions/site
5.27 substitutions/site

Appendix 2.1

VHSv sequences per strain, isolate, host species, locality, GenBank accession numbers, and references. Taxon names in quotation marks are questionable or have been reclassified in this work. *=unique phylogenetic isolates used in our trees. Dash mark (-) under latitude and longitude indicate that specific localities could not be obtained.

Gene	VHSv Type	Host Species	Isolate	Accession No.	Homologous Sequences: Isolate (Accession No.)	Reference(s)	Locality	Latitude	Longitude
G	**"I"/ FiA**	*Oncorhynchus mykiss*	*FiA06.02, *FiA10.02, *FiA17.02, *FiA19.04, *FiP03.00	AM086367, AM086371, AM086378, AM086383, AM086357	—	Raja-Halli et al., (2006)	FIN	—	—
		" "	FiA12.02	AM086373	FiA08.02(AM086369)	" "	" "	—	—
	"I"	*Esox lucius*	X4(1999)	AJ487055	—	Thiéry et al., (unpub.)	Moselle, FRA	49.008	6.583
		Oncorhynchus mykiss	*Hededam	Z93412	—	Stone et al., (1997)	Spjarup Hededamme, DNK	56.648	9.271
		" "	R104(1993), S1(1994)	AJ487078-79	—	Thiéry et al., (unpub.)	Pas-de-Calais, FRA	50.660	2.910
		" "	*DK-F1	AF345857	—	Einer-Jensen et al., (2004)	Kølkær Dambrug, DNK	56.093	9.110
		" "	J59(1985)	AJ487077	—	Thiéry et al., (unpub.)	Gironde, FRA	44.650	1.270
		" "	1458(1990)	AJ487075	—	" "	Finistere, FRA	47.850	3.930
		" "	XP4(1991)	AJ487074	—	" "	Isere, FRA	45.050	5.050
		" "	R94(1993)	AJ487073	—	" "	Meuse, FRA	48.690	5.550
		" "	F48(1981)	AJ487072	—	" "	Finistere, FRA	47.850	3.930
		" "	S93(1994), T6(1995), T7(1995)	AJ487069-71	—	" "	Seine-maritime, FRA	49.760	0.363
		" "	R22(1993), R73(1993)	AJ487067-68	—	" "	Pas-de-Calais, FRA	50.660	2.910

	“ “	W94(1998)	AJ487066	__	“ “	Jura, FRA	46.580	5.877
	“ “	YP41(1992)	AJ487065	__	“ “	Doubs, FRA	46.883	5.016
	“ “	S41(1994)	AJ487064	__	“ “	Eure-et-Loir, FRA	48.311	1.252
	“ “	5947(1991)	AJ487063	__	“ “	Pyrenees-Atlantiques, FRA	48.311	1.252
	“ “	S79(1994)	AJ487062	__	“ “	Somme, FRA	49.840	1.890
	“ “	5796(1993)	AJ487061	__	“ “	Dordogne, FRA	45.033	0.600
	“ “	3771(1990)	AJ487060	__	“ “	Landes, FRA	47.867	5.333
	“ “	4957(1994), 5785(1993)	AJ487056-57	__	“ “	Dordogne, FRA	45.033	0.600
	“ “	R68(1993)	AJ487054	__	“ “	Orne, FRA	48.433	0.082
	“ “	XP45(1991)	AJ487053	__	“ “	Jura, FRA	46.580	5.877
	“ “	R27(1993)	AJ487052	__	“ “	Finistere, FRA	48.000	3.930
	“ “	XP68(1991)	AJ487051	__	“ “	Puy-de-Dome, FRA	46.000	2.000
	“ “	957	Z93420	__	Stone et al., (1997)	River mass, Noord-Brabant, NLD	51.770	5.539
	“ “	Grasmuck	Z93419	__	“ “	Grasmuck, FRA	33.780	96.680
	“ “	670	Z93413	__	“ “	Amersfoort, Utrecht, NLD	51.770	5.539
	“ “	609	Z93417	__	“ “	Limburg, NLD	50.917	5.767
	“ “	448	Z93418	__	“ “	River mass, Noord-Brabant, NLD	51.770	5.539
	“ “	Rindsholm	Z93416	__	“ “	Hededam, DNK	51.772	5.539
	“ “	83-53 (England), F1	Z93411, Z93408	__	Lumsden et al., (2007)	DNK	54.000	2.000
	“ “	S45(1994)	AJ487059	R67(1993)(AJ487058)	Thiéry et al., (unpub.)	Aveyron, FRA	44.618	2.574
	Salmo trutta	Hededam Denmark 70, 23.88 France 1988	U28798, U88054	__	Benmansour et al., (1997)	Spjarup Hededamme, DNK, FRA	56.648	9.271
	“ “	J77(1981)	AJ487076	__	Thiéry et al., (unpub.)	Ille-et-Villaine, FRA	47.760	1.980
	Scophthalmus maximus	7321(5927)	Z93409	__	Stone et al., (1997)	Tuebingen, DEU	51.060	9.280
	__	A10182	A10182	__	__	__	__	__
Ia	*Esox lucius*	Dsa16257-05	EU708789	Dsa14893-05(EU708787)	Enzmann et al., (unpub.)	Tuebingen, DEU	10.180	10.180
	Onchorhychus mykiss	Fil3	NC000855	Fil3(Y1826), Dfi13-83(EU708759), DK-3946(AY546586)	Schutze et al., (1999)	Baltic S., DEU	51.000	10.000

“	“	French strain 07-71	AJ233396	FR-0771(AY546616, X59148)	Xing et al., (unpub.), Einer-Jensen et al., (2004), Thiry et al., (1991)	Seine-maritime, FRA	49.760	0.363
“	“	80-96	AY612120	299-94Austria (AY612110)	Schachner et al., (unpub.)	AUS	__	__
“	“	PL14/05	DQ897664	PL11/05, PL30/05, PL20/05, PL22/05 (DQ897660-DQ897663)	Chretiennot-Dinet et al., (unpub.)	POL	__	__
“	“	UK-J167	JN180851	__	Stone et al., (2008)	UK	__	__
“	“	189-98	AY612105	65-98 (AY612117)	“ “	AUS	__	__
“	“	*Dwb97-04	EU708816	Dwb88-04(EU708815), Dwb48-03(EU708813), Dstg84-03(EU708802)	Enzmann, (unpub.)	Baden Wuerttemberg, DEU	48.800	9.190
“	“	*DK-200079-1	AY546613	DK-200051-1(AY546611), DK-9995361(AY546604), DK-20079-1(AY546603)	Einer-Jensen et al., 2004	Nærrild Dambrug, DNK	55.840	8.490
“	“	29-95	AY612109	22-94(AY612106)	Schachner et al., (unpub.)	AUS	__	__
“	“	303-94	AY612111	__	“ “	“ “	__	__
“	“	*DK-5741	AY546591	DK-740(AY546590), DK-5727(AY546589)	Einer-Jensen et al., (2004)	Nærrild Dambrug, DNK	55.840	8.490
“	“	*FiA03.04	AM086382	FiA04.02(AM086366), FiA01a_b.00(AM086354), FiA03.03(AM086379), FiA15.02(AM086376), FiA07.02(AM086368), FiA03.02(AM086365), FiA04.01(AM086361), FiU01.03(AM086381), FiA02b.01(AM086359), FiA09.02(AM086370), FiA03a_b.01(AM086360), FiA02a.01 (AM086358), FiP01a_c.00(AM086355), FiA16.02(AM086377), FiA18.03(AM086380), FiA05.01(AM086362)	Raja-Halli et al., (2006)	FIN	__	__
“	“	*Dau24-02	EU708737	Dstg49-03(EU708799), Dstg38-03(EU708797), Dstg38-03(EU708795), Dstg74-03(EU708801), Dstg42-06(EU708808), Dns166-06(EU708770)	Enzmann et al., (unpub.)	DEU	__	__

“	“	*Dstg50-2-04	EU708800	Dstg37-04(EU708796)	“	“	—	—
“	“	*Dsh-fv2-07	EU708792	Dns7-07(EU708768), Dns6-07(EU708767)	“	“	—	—
“	“	*Dau1124-00	EU708749	Dau519-00(EU708746), Dau463-00(EU708745)	“	“	—	—
“	“	*FiA11.02	AM086372	FiA13.02(AM086374), FiA14.02(AM086375)	Raja-Halli et al., (2006)	FIN	—	—
“	“	*FiP03.01	AM086364	FiP03.00(AM086357), FiP04.01(AM086363), FiP02a_b.00(AM086356)	“	“	—	—
“	“	*Au8/95, 11-96, 101-97, 152-97, 13-95, 159-98, 169-94, 17-95, 28-95, 38-97, *42-96, 50-95, 58-95, *74-98, 83-94, *08-95, *94-98	AY546570, AY612096, AY612098-02, AY612104, AY612108, AY612112-15, AY612118, AY612121-23	—	Schachner et al., (unpub.)	AUS	—	—
“	“	*Au62-96	AY612116	78-96(AY612119),124-98(AY612097), 286-94 (AY612107)	“	“	—	—
“	“	*DK-200149	AY546607	DK-200148 (AY546606), DK-200070-4 (AY546612), DK-200020-3, DK-200027-3, DK-200029-1 (AY546608- AY546610)	“	DNK	—	—
“	“	*CH-FI262BFH, DK-3971, *DK-6045, *DK-6137, *DK-7380, *DK-7974, *DK-9595168, *DK-9695377, *DK-9795568, DK-9895024, *DK-9895093, *DK-9995007, *DK-9995144, *DK-9895174, *DK-200098, *DK-5151	AY546571, AY546587, AY546592-02, AY546605, AF345859	—	“	“	—	—
“	“	Ckc-1-5, Bolu/06	JF415086-91	—	Isidan et al., (unpub.)	TUR	—	—
“	“	*Dau42-99, *Dau56-02, *Dau170-04, *Dau201-05, *Dau412-02, *Dau543-03, *Dau1503-01, *Dau1556-98, *Daulab,	EU708739, EU708741-44, EU708747, EU708751-53, EU708757-58, EU708760-61,	—	Enzmann et al., (unpub.)	Saxony, DEU	51.050	13.750

		*Db493-08, *Db905-03, *Dfr195-05, *Dfr404-05, *Dfr2247-06, *Dfr2868-06, *Dns46-07, *Dns13462-03, *Dri07-05, *Dri12-95, *Dsa12-99, *Dsa19-7-93, *Dsa90-98, Dsa67-01, *Dsa141-7-94, *Dsa142-96, *Dsa6915-02, *Dsa14894-05, *Dsa21583-03, *Dsan152, *Dstg22-03, *Dstg31-07, *Dstg36-06, *Dvgeig, *Dwb86-04	EU708764-65, EU708769, EU708771, EU708773-74, EU708779-81, EU708783-86, EU708788, EU708790-91, EU708794, EU708806-07, EU708812, EU708814						
“	“	Dstg89-1-07	EU708811	Dstg82-1-07(EU708810), Dri47-06(EU708776)	“	“	“	“	“	“
“	“	*Dstg54-1-07	EU708809	Dstg21-07(EU708804)	“	“	“	“	“	“
“	“	*Dstg28-06	EU708805	Dstg8-06(EU708803)	“	“	“	“	“	“
“	“	*Dstg44-04	EU708798	Dfr148-07(EU708762), Dfr378-07(EU708763), Dri04-08(EU708775)	“	“	“	“	“	“
“	“	*Dsa82-99	EU708782	Dsa9-99(EU708778), Dsa05-00(EU708777), Dri47-06(EU708776), Dau1402-97(EU708750), Dau686-97(EU708748)	“	“	“	“	“	“
“	“	*Db416-92	EU708756	Db359-92(EU708755)	“	“	“	“	“	“
“	“	*Dau10-97	EU708736	Dau9-97(EU708735)	“	“	“	“	“	“
“	“	*Au28-95, *Au62-96, *Au77-99, *Au299-94, *Au917-04, *Datt107	EU708729-34	__	“	“	AUS		__	__
“	“	FA-25	EU336985	__	Duesund et al., (2010)		Moere og Romsdal, NOR	62.310	6.900	
“	“	*1455/07, SLO 3, SLO 1, SLO 5-7	GQ292534, GQ153531-34	__	Toplak et al., (unpub.)		SLO		__	__
“	“	*Strain 14-58	AF143863	__	Betts and Stone (2000)		FRA		__	__
“	“	*Daurxs-98, *Dsteinbutt	EU708754, EU708793	__	Enzmann et al., (unpub.)		DEU		__	__
“	“	Dau52-98	EU708740	Dau35-97 (EU708738)	“	“	“	“	__	__
“	“	TR-Bs13/15H	AB231161	__	Nishizawa et al., (2006)		Trabzon, TUR	40.870	40.380	

	Salmo trutta	02-84	Z93404	U28800	" "	Seine-maritime, FRA	49.760	0.363
	" "	*Strain 23-75	FN665788	FR-2375(AY546617), strain 23-75 partial(Z93415)	Biacchesi et al., (2010)	FRA	__	__
	" "	23-75	U28799	__	Benmansour et al., (1997)	" "	__	__
	Thymallus arcticus	*Dri01-06	EU708772	__	Enzmann et al., (unpub.)	Baltic S., DEU	51.000	10.000
	" "	173-98	AY612103	__	Schachner et al., (unpub.)	AUS	__	__
Ib	*Clupea pallasii*	*strain-96-43	AF143862	__	Betts and Stone (2000)	English Channel, GBR	50.600	0.000
	" "	*UK-MLA98/6HE1	AY546631	SE-SVA32(AY546627), SE-SVA30(AY546625)	Einer-Jensen et al., (2004)	North S., DNK; Kattegat, DNK	56.000, 56.320	10.000, 11.560
	" "	*DK-1p8	AY546573	DK-e62(AY546572), DK-1p12(AY546574), Dglasaal(EU708766)	" "	Baltic S.	55.951	17.767
	" "	*SE-SVA31	AY546626	SE-SVA29(AY546624), DK-6p403(AY546584)	" "	Kattegat, DNK	56.320	11.560
	" "	*NO-F/2009	HM632035	__	Gjerset et al., (unpub.)	Finnmark, NOR	70.619	24.621
	Gadus morhua	*Cod Ulcus	Z93414	__	Stone et al., (1997)	Little belt, Baltic S.	55.500	9.650
	Gaidropsarus mediterraneus	*DK-1p40	AY546575	__	Einer-Jensen et al., (2004)	Baltic S., DEU	51.000	10.000
	Limanda limanda	*DK-5e59	AY546583	__	Campbell et al., (2009)	Kattegat, DNK	56.320	11.560
	Onchorhychus mykiss	*SE-SVA1033	AY546623	FJ460591, DK-4p37(FJ460590), DK-4p37(AY546580)	" "	" "	" "	" "
	" "	*SE-SVA-14	AY546622	__	" "	" "	" "	" "
	" "	*CH150208	FJ384761	__	Duesund et al., (2010)	Storfjorden Møre and Romsdal, NOR	62.360	6.160
	Paralichthys olivaceus	*JP96KRRV9601	DQ401190	__	Elsayed et al., (2006)	Kagawa Prefecture, JPN	34.340	134.043
	Sprattus sprattus	*DK-1p86	AY546579	__	" "	Baltic S.	54.834	19.200
"Ic"/ Ia	*Onchorhychus mykiss*	*DK-5123	AY546588	__	Einer-Jensen et al., (2004)	Vejen lille Vandmølle Dambrug, DNK	55.484	9.151
	" "	*DK-5131, *DK-2835	AF345858, AY546585	__	" "	Klapmolle Damburg, DNK	56.078	8.351
	""	Klapmolle	Z93410	__	Stone et al., (1997)	" "	" "	" "
Id	" "	*NO-A16368G	AY546621	__	Einer-Jensen et al., (2004)	Vestrefjord, NOR	62.617	6.660
	" "	FiP02a_b.00	AM086356	F1-ka422(AY546615)	Raja-Halli et al., (2006)	Pyhtaa kunta, FIN	62.000	26.000
	" "	F1-Ka66	AY546614	__	Einer-Jensen et al., (2004)	Gulf of Bothnia, FIN	60.000	24.000

Ie	" "	GE-1.2	AY546619	__	" "	GEO	41.780	44.880
	Scophthalmus maximus	TR-SW13G, TR-Bs13115H	AB231160, Ab231161	__	Nishizawa et al., (2006)	Trabzon, TUR	41.020	39.750
II	*Clupea harengus*	*DK-1p53	AY546577	__	Einer-Jensen et al., (2004)	Baltic S.	54.834	19.200
	" "	ka364_04, ka558_04, ka565_04, ka655_06, *ka350_06, ka366_04, ka371_04, ka388_04, ka392_04, *ka396_04, ka427_04, *ka436_04, *ka494_05, *ka560_04, *ka564_04, ka575_04, *ka664_04	HQ112199, HQ112225, HQ112232-35, HQ112237-45, HQ112248	__	Gadd et al., (2011)	Archipelago S., FIN	60.290	21.290
	" "	ka383_04	HQ112236	ka363_04(HQ112198), ka381_04(HQ112205-07), ka367_04(HQ112201)	" "	" "	" "	" "
	" "	ka365_04	HQ112200	ka368_04; ka369_04; ka380_04(HQ112202-04), ka385_04; ka387_04; ka389_04; ka390_04; ka391_04, ka393_04-ka395_04; ka397_04; ka414_04; ka428_04; ka557_04; ka558_04; ka559_04; 561_04-ka563_04; ka565_04; ka570_04; ka574_04 (HQ112208-21, HQ112222-24, HQ112226-27)	" "	" "	" "	" "
	" "	*ka646_04	HQ112246	ka645_04(HQ112228)	" "	" "	" "	" "
	" "	*ka663_06	HQ112247	ka664_06(HQ112231), ka662_06(HQ112230)	" "	" "	" "	" "
	Lampetra fluviatilis	*F1-lamprey 739.03	GQ504014	F1-lamprey 743.03 (GQ504013)	Gadd et al., (2010)	Lestijoki R., FIN	64.062	23.657
	Sprattus sprattus	*DK-1p55	AY546578	DK-1p52(AY546576)	Einer-Jensen et al., (2004)	Baltic S.	54.834	19.200
III	*Anguilla anguilla*	*FR-L59x(1987)	AJ487080	AY546618	Thiéry et al., (2002)	Loire-Atantique, FRA	47.430	1.660
	Clupea harengus	CH18.03.08, CH150208	GU066860, FJ384761	__	Duesund et al., (2010)	Storfjorden Møre and Romsdal, NOR	62.360	6.160
	Clupea pallasii	*DK-4p168	AY546582	__	Einer-Jensen et al., (2004)	Skagerrak, North S.	57.985	10.601
	Gadus morhua	*UK-H17/5/93, H191, H17/5	AY546630, Z93406-07	__	Stone et al., (1997)	E. Shetland, North S., UK	60.283	0.044

	Merlangius merlangus	*DK-4p101	AY546581	__	Einer-Jensen et al., (2004)	North S.	57.433	7.550
	Melanogrammus aeglefinus	*UK-H17/2/95	AY546629	__	" "	E. Shetland, North S., UK	60.283	0.044
	Onchorhychus mykiss	BV210208-1, H050308-2, KH230208-2, KV010308-4, NO/650/07	FJ362511-14, AM920657	__	Duesund et al., (2010), Dale et al., (2009)	Storfjorden Møre and Romsdal, NOR	62.360	6.160
	" "	*FA281107	EU481506	V230308-5(FJ362515), BV060408-52(FJ362510), NO-2007-50-385 (EU547740), FA28.02.08V(GU121101), FA280208-SG(GU121099)	Duesund et al., (2010)	" "	" "	" "
	S. trutta	17-91	Z93431	__	Stone et al., (1997)	FRA	46.000	2.000
	Scophthalmus maximus	814, *IR-F13.02.97, *UK-860/94, Scotland 1995	Z93405, AY546620, AY546628, U88056	__	Stone et al., (1997), Benmansour et al., (1997)	Gigha, W. SCO	55.670	5.800
	Trisopterus esmarki	*UK-MLA98/6PT11	AY546632	__	Einer-Jensen et al., (2004)	Beryl, North S.	59.380	1.450
IVa	*Acanthopagrus australis*	YBS05	GU265815	LH03(GU265814), KF05(GU265813), BF05-2(GU265810), BF05-1(GU265809)	Oh et al., (unpub.)	S. KOR	33.000	127.500
	Brachymystax lenok	Inje07	FJ561738	__	Jeon et al., (unpub.)	" "	" "	" "
	" "	Inje07	FJ561738	__	" "	" "	" "	" "
	Cymatogaster aggregata	BC-sp-02	DQ473301	__	Hedrick et al., (2003)	British Columbia, NW Vancouver Is., CAN	49.233	-123.100
	Clupea harengus	*ME03	DQ401192	__	Elsayed et al., (2006)	N. Pacific, USA	47.622	-122.638
	C. pallasii	EB#7	Z93428	__	Stone et al., (1997)	Elliott Bay, WA, USA	" "	" "
	" "	BC'93	Z93427	__	" "	Prince Rupert Sound, BC, CAN	54.312	-130.327
	" "	Campbell River, BC 1993	U88051	__	Benmansour et al., (1997)	Campbell River, BC, CAN	50.024	-125.248
	" "	WA-93, Elliot Bay, WA 1993,	DQ473303, U88055	__	Hedrick et al., (2003), Benmansour et al., (1997)	Elliott Bay, WA, USA	47.622	-122.638
	" "	AK93#1, Prince William Sound AK, 1990	Z93430, U88052-53	__	Stone et al., (1997), Benmansour et al., (1997)	Prince William Sound, AK, USA	61.622	-146.863
	" "	*BC93372, *BC99010,	DQ401186,	__	Elsayed et al., (2006)	Beaver Cove, BC, CAN	50.591	-126.899

		*BC99001	DQ401194-95					
	Gadus macrocephalus	AK'93	Z93429	__	Stone et al., (1997)	Prince William Sound, AK, USA	61.622	-146.863
	" "	NA-7	Z93425	__	" "	" "	" "	" "
	" "	*AK-90 and Clearwater WA, 1991	DQ473302, U88050	__	Hedrick et al., (2003), Benmansour et al., (1997)	" "	" "	" "
	" "	NA-6	Z93424	__	Stone et al., (1997)	" "	" "	" "
	Onchorhychus kisutch	*Makah, Elok, NA-5, NA-8	U28747, Z93421-23, Z93426	__	" "	Clearwater and Bogachiel R., WA, USA	48.289, 47.929	-124.652, -123.837
	O. mykiss	RtGw5	HQ687073	RtGw8(HQ687072)	Suebsing et al., (unpub.)	S. KOR	33.000	127.500
	" "	RtGw11, RtGw10	HQ687070-71	__	" "	" "	" "	" "
	Paralichthys olivaceus	FJeju05	FJ811902	FYeosu05(FJ811901), FWando05(FJ811900), FYG08(GU265812)	Kim and Park (unpub.)	" "	" "	" "
	" "	KRRV9822	AB179621	__	Byon et al., (2006)	JPN	__	__
	" "	FWando08	GU265811	__	Oh et al., (unpub.)	S. KOR	33.000	127.500
	" "	*AY167587	AY167587	__	Kim and Park (unpub.)	" "	" "	" "
	" "	OfGn	HQ687076	PhGn(HQ687075), PcGn(HQ687074)	Suebsing et al., (unpub.)	" "	" "	" "
	" "	*JP99Obama25, 9601, JF00Ehi1, #25	DQ401191, AB060727, AB490792, AB060725	__	Elsayed et al., (2006), Ito (unpub.), Nishizawa (unpub.)	Kagawa Prefecture, Ehime and Wakasa Bay JPN	34.340, 33.750, 35.200	134.043, 132.600, 134.240
	" "	*KJ2008	JF792424	__	Kim and Kim (unpub.)	S. KOR	33.000	127.500
	Salmo salar	*BC98250, *BC99292	DQ401187-88	__	" "	BC, CAN	__	__
	" "	*WA91 Clearwater	DQ401189	__	" "	WA, USA	47.622	-122.638
	Sardinops sagax	*CMLs02	DQ473300	__	Hedrick et al., (2003)	Moss Landing, CA, USA	36.804	-121.786
	" "	*CMAs02	DQ473299	__	" "	Malibu, CA, USA	34.030	-118.780
	" "	*CS01	DQ473298	__	" "	Los Angeles, CA, USA	33.780	-118.050
	" "	BC-s-99	DQ473296	__	" "	Queen Charlotte St., BC, CAN	54.214	-130.693
	Thaleichthys pacificus	*Oe01	DQ473297	__	" "	Sandy R., OR, USA	45.568	-122.400
IVb	*Ambloplites rupestris*	*TAVgr08-03 (vcG009)	HQ623441	__	" "	North Point Marina, L. MI, USA	42.499	-90.689

Aplodinotus grunniens	*U13653-1 (vcG002)	HQ453209	RG06(EF564588), U13653-2, TAVgr06-02, TAVgr06-27, TAVgr08-04, TAVgr06-19, TAVgr06-03, TAVgr06-04, TAVgr06-31, TAVgr06-32, TAVgr06-33, TAVgr06-34, TAVgr06-35, TAVgr06-36, TAVgr06-39, TAVgr06-37, TAVgr06-38, TAVgr06-24, TAVgr06-20, TAVgr06-25, TAVgr06-21, TAVgr06-26, TAVgr06-23, TAVgr06-40, TAVgr06-22, TAVgr08-07, TAVgr07-08, TAVgr08-08, TAVgr08-02, TAVgr08-06, TAVgr08-05, TAVgr06-30, TAVgr06-01, TAVgr07-18, TAVgr07-19	Thompson et al., (2011)	Bay of Quinte, L. Ont., CAN; St. Lawrence R., NY, USA; Dunkirk Harb., L. Erie, NY, USA; L. Huron, USA; L. Michigan, WI, USA; Fairhaven State Park, L. Ont., NY, USA; Oswego, L. Ont., NY, USA; Tibbetts Ck., L. Ont. USA; W. of Rochester, L. Ont., USA; L. Skaneateles, NY, USA	43.968, 44.250, 44.327, 44.127, 44.257, 44.323, 44.248, 44.254, 44.187, 44.175, 44.323, 44.172, 44.323, 44.242, 44.187, 44.268, 44.254, 43.340, 42.490, 45.625, 42.799, 43.350, 43.450, 44.116 43.216 42.490	-77.629, -75.000, -75.937, -76.333, -76.134, -76.014, -76.014, -76.014, -76.014, -76.964, -76.014, -76.247, -75.935, -76.098, -76.225, -76.014, -76.150, -75,910, -79.338, -84.468, -87.760, -76.690, -76.510 -76.333, -77.633, -79.338
Cyprinus carpio	*TAVgr07-12 (vcG003)	HQ623435	TAVgr07-17, TAVgr07-13	Thompson et al., (2011)	Dunkirk Harb., L. Erie, NY, USA; Little Salmon R., NY, USA; Skaneateles L., NY, USA	42.490, 43.459, 42.950	-79.338, -76.228, -76.230
Esox masquinongy	*MI03GL	GQ385941	DQ401193, TAVgr05-01, 0601FD, 0602SB, 0603BG, TAVgr06-48, TAVgr06-51, TAVgr06-52, TAVgr06-50, TAVgr06-47, TAVgr06-49, TAVgr09-03, TAVgr09-04, TAVgr09-10, TAVgr09-11, TAVgr07-09, TAVgr09-05,	Elsayed et al., (2006), Thompson et al., (2011)	L. St. Clair, MI, USA; Dunkirk Harb., L. Erie, NY, USA; L. Erie, USA; L. Erie, OH, USA Sandusky Bay, L. Erie, OH, USA; W. L. Erie, OH, USA;	42.390, 42.343, 42,634, 42.631, 42.615, 42.475, 42.490, 41.755,	-82.911, -82,902, -82.777, -82,765, -82,757, -82.879, -79.338, -81.286,

			TAVgr06-07, TAVgr06-08, TAVgr06-09, TAVgr06-10, TAVgr06-11, TAVgr06-15, TAVgr06-53, TAVgr07-21, TAVgr07-22, TAVgr08-09, TAVgr08-10, TAVgr08-11, TAVgr06-05, TAVgr06-06, TAVgr06-12, TAVgr06-13, TAVgr06-14, TAVgr06-17, TAVgr06-18, TAVgr06-46, TAVgr06-44, TAVgr06-43, TAVgr06-45, TAVgr09-01, TAVgr09-02, TAVgr09-12, OMNR 5577, OMNR 5583, OMNR#5579, TAVgr06-28, TAVgr07-06, TAVgr07-07, TAVgr06-29, TAVgr10-01, TAVgr09-09, TAVgr07-02, TAVgr07-03, TAVgr07-14, TAVgr07-15, TAVgr07-16, TAVgr07-10, TAVgr07-11, TAVgr09-32, TAVgr09-33, TAVgr07-05, TAVgr09-13		Cheboygan Bay, L. Huron, MI, USA; Swan R., L. Huron, MI, USA; Thunder Bay, L. Huron, MI, USA; L. Mich. MI, USA; Sturgeon Bay, L. Mich., WI, USA; Hamilton Harb. L. Ont., ON, CAN; Thames R., L. ON, ON, CAN; Irondequoit Bay, L. Ont., NY, USA; L. Ont., Rochester, NY, USA; Sodus Bay, L. Ont., NY, USA; Apostle Is., L. Super., WI, USA; Baseline L., MI, USA; Budd L., MI, USA; Cayuga-Seneca Canal, NY, USA; Ransomville, NY, USA; Clear Fork Res., OH, USA; L. Winnebago, WI, USA; Oak Ck./Grant Pk., WI, USA	41.474, 41.492, 45.718, 45.502, 45.050, 43.600, 44.860, 44.514, 43.295, 42.328, 43.200, 43.233, 42.257, 47.085, 42.427, 44.015, 42.910, 43.240, 40.716, 44.028, 42.926	-82.703, -82.667, -84.374, -83.783, -83.200, -86,916, -87,393, -87.831, -79.772, -82,472, -77.526, -76.650, -76.966, -90.641, -83.899, -84.788, -76.910, -79.920, -82.643, -88.421, -87.770
Lepomis gibbosus	*TAVgr07-04 (vcG006)	HQ623438	__	“ “	Budd L., MI, USA	44.015	-84.788
L. macrochirus	*TAVgr07-01 (vcG005)	HQ623437	__	“ “	“ “	“ “	“ “
“ “	*TAVgr07-20 (vcG007)	HQ623439	__	“ “	East Harb., L. Erie, OH, USA	41.541	-82.789
Micropterus dolomieu	*TAVgr07-24 (vcG008)	HQ623440	__	“ “	Sturgeon Bay, L. Michigan, MI, USA	44.884	-87.338
Morone chrysops	*TAVgr06-16 (vcG004)	HQ623436	TAVgr08-01	“ “	Western Basin, L. Erie, OH, USA; Clear Fork	41.492, 40.716	-81.667, -82.643

							Res., OH, USA		
		Perca flavescens	*TAVgr09-17 (vcG010)	HQ623442	__	" "	L. Michigan, WI, USA	43.039	-87.802
		Percopsis omiscomaycus	*TAVgr06-53 (vcG011)	HQ623443	__	" "	Central Basin, L. Erie, OH, USA	41.755	-81.286
	IVc	*Fundulus heteroclitus*	*CA-NB00-01	EF079896	__	Gagné et al., (2007)	Ruisseau George Collette, near Bouctouche, NB, CAN	46.450	-64.682
		Gasterosteus aculeatus	*CA-NB00-02	HQ168405	__	" "	" "	" "	" "
		Morone saxatilis	*CA-NB04-01b	HQ453208	__	" "	Miramichi Bay, Baie du Vin, NB, CAN	47.163	-64.574
		" "	*CA-NB02-01	EF079897	CA-NB04-01(EF079898)	" "	" "	" "	" "
		S. trutta	*CA-NS04-01	EF079899	__	" "	French R., NS, CAN	45.576	-62.425
N	**I**	*Onchorhychus kisutch*	*Hededam	Z93412, EU932929	__	Stone et al., (1997)	Spjarup Hededamme, DNK	56.648	9.271
		O. mykiss	DK-F1	AY356633	__	Snow et al., (2004)	" "	" "	" "
	Ia	*O. mykiss*	*Strain 14-58	AF143863	__	Betts and Stone (2000)	FRA	__	__
		" "	DK-3956	AY356641	DK-3955(AY356640)	Snow et al., (2004)	DNK	__	__
		" "	AU-8/95	DQ159192	__	Einer-Jenesen et al., (unpub.)	AUS	__	__
		" "	DK-6137	DQ159190	__	" "	Nærrild Dambrug, DNK	55.840	8.490
		" "	DK-7843	AY356650	DK-7217, DK-7300, DK-7309, DK-7655 (AY356646-AY356649), DK-795265(AY356637), DK-9795159(AY356636)	Snow et al., (2004)	Vingsted Dambrug, DNK	55.670	9.380
		" "	DK-6143, DK-5133, DK-9895174	AY356645, AY356642, AY356638	__	" "	Nærrild Dambrug, DNK	55.840	8.490
		" "	*Fil3	NC_000855	DK-6047(AY356644), DK-5243(AY356643), Fil3(Y18263), DK-3345(AY356639)	" "	Baltic S., DEU	50.000	10.000
		" "	* X73873	X73873	__	Bernard et al., (1990)	" "	" "	" "
		" "	DK-9695297	AY356635	DK-9695152(AY356634)	Snow et al., (2004)	DNK	__	__
		" "	*French0771	AJ233396	pathogenic 07-71(D00687)	Xing et al., (unpub.)	Seine-Maritime, FRA	49.500	0.500
		" "	*DK3592-B	AF012093	DK-5243(AY356643)	Einer-Jensen et al., (unpub.)	Voldbjerg Dambrug, DNK	55.200	11.200
		Salmo trutta	*FR23-75	FN665788	FR-2375(DQ159191)	Biacchesi et al., (2010)	Eure, FRA	49.800	1.150

	Scophthalmus maximus	7321(5927)	AJ130922	__	Snow et al., (1999)	DEU	__	__
Ib	*Clupea pallasii*	5p201	AY356676	5p213(AY356677), 5p31(AY356674)	Snow et al., (2004)	Kattegat, DNK	56.320	11.560
	" "	NO-F/2009	HM632036	__	Gjerset et al., (unpub.)	Finnmark, NOR	70.619	24.621
	" "	*DK-1p8	GQ325430	96-43(AY356667)	Campbell et al., (2009), Snow et al., (2004)	North S., DNK	56.000	10.000
	" "	*strain-96-43	AF143862	5p680(AY356713), 1p50(AJ130920), 1e62(AY356668), 5p48(AY356675), 1p3, 1p8, 1p12, 1p64, 1p85 (AY356651- AY356655), 1p93, 1p109-11, 1p116, 1p120-1, 1p124-5, 1p128, (AY356657-AY356666), 5e62-3, 5p11, 5p26 (AY356670- AY356673), 5p251, 5p253, 5p263, 5p276, 5p393, 5p405, 5p439, 5p441-2, 5p454 (AY356678- AY356687)	Betts and Stone (2000), Snow et al., (2004), Snow et al., (1999)	English Channel, GBR	50.600	0.000
	" "	*DK-5e59	AY356669 DK-5e59 (GQ325429)	7e560(AY356718)	Snow et al., (2004), Campbell et al., (2009)	Kattegat, DNK	56.320	11.560
	" "	*CH180308-N	GU066860	__	Duesund et al., (2010)	NOR	__	__
	Enchelyopus cimbrius	1p40	AJ130919	__	Snow et al., (1999)	Baltic S.	57.118	19.686
	Gadus morhua	*Cod Ulcus	Z93414	DK-M.Rhabdo(AY356632)	Stone et al., (1997), Snow et al., (2004)	Little belt, Baltic S.	55.500	9.650
	Micromesistius poutassou	*DK-4p37	FJ460590	__	Campbell et al., (2009)	North S., DNK	57.951	5.267
	" "	DK-9695297	AY356635	DK-9695152(AY356634)	Snow et al., (2004)	" "	" "	" "
	" "	*SE-SVA-1033	FJ460591	__	Campbell et al., (2009)	W./E. 1255009 S./N. 6410709, SWE	57.750	11.680
	" "	*SE-SVA-14	GQ325428	__	Campbell et al., (2009)	SWE	__	__
	Onchorhychus mykiss	UKMLA98/6HE1	GQ325431	__	" "	UK	__	__
	Sprattus sprattus	1p50	AJ130920	__	Snow et al., (1999)	Baltic S.	57.118	19.686
	" "	1p86	AY356657	__	Snow et al., (2004)	Baltic S., DEU	56.500	19.000
	" "	DK-1p52	AY356744	__	" "	" "	" "	" "

Ie	*Onchorhychus mykiss*	GE-1.2	DQ159189	__	Einer-Jensen et al., (unpub.)	GEO	41.680	44.880
II	*Clupea harengus*	5p557	AY356691	5p551(AY356690)	Snow et al., (2004)	Baltic S., DEU	56.500	19.000
	" "	1p52	AY356743	1p54, *f*1p49, 1p55(AY356744-AY356746), 1p53(AJ130921)	" "	" "	" "	" "
III	*Argentina sphyraena*	4p51	AY356736	__	" "	North S.	57.951	5.267
	Clupea harengus	5p448	AY356688	5p449(AY356689), 5p457, 5p508, 5p513, 5p514, 5p524, 5p527, 5p547, 5p663, 5p664, 5p665, 5p666, 5p669, 5p670, 5p671, 5p673, 5p674, 5p675, 5p676, 5p677, 5p679, 5p678, 5p680, 7p97, 7p37, 7e538, 7e544 (AY356692- AY356717)	" "	Baltic S., DEU	56.500	19.000
	Gadus morhua	H16/7/95	AJ130923	4PT2-3(AY356737-8), 4PT5-6, 4PC1 (AY356740-AY356742), 4P168, 6PT11-12(AY356732-4), 6PT15(AY356732), 6PT1, 6PT16, 6PT10, 6PT17, 6PT7-8, H17/5/93, H19/1/95(AY356721-AY356728)	Snow et al., (1999)	North S.	57.951	5.267
	" "	CO534	AY356735	__	Snow et al., (2004)	" "	" "	" "
	Melanogrammus aeglefinus	H17/2/95	AJ130924	H17/1/95(AY356720)	Snow et al., (1999)	North S., GBR	60.700	1.650
	" "	2p51	AJ130917	__	Snow et al., (unpub.)	Skagerrak, North S.	57.985	10.601
	Merlangius merlangus	4p101	AJ130918	__	Snow et al., (1999)	Skagerrak, North S.	57.985	10.601
	" "	6WH1	AY356729	__	Snow et al., (2004)	North S.	57.951	5.267
	Onchorhychus mykiss	*FA281107	EU481506	BV060408-52(FJ362510), KV010308-4(FJ362514), FS280208-V(GU121102), V230308-5(FJ362515)	Duesund et al., (2010)	NOR	__	__
	" "	*FA280208-S	GU121100	__	" "	" "	__	__
	Reinhardtius hippoglossoides	GH48	AJ849490	GH30, GH32-37, GH40, GH42-43, GH45-47 (AJ849477- AJ849489)	López-Vázquez et al., (unpub.)	Flemish Cap, ESP	__	__

		Scrophthalmus maximus	F 13.02.97	AJ130916	__	Snow et al., (1999)	IRL	__	__
		" "	860/94	AJ130915	__	" "	Gigha, W. SCO	55.670	5.800
		Trisopterus esmarki	4PT4, 6PT14, 4PT1	AY356739, AY356730, AY356719	__	Snow et al., (2004)	North S.	57.951	5.267
	IVa	*Clupea pallasii pallasii*	PhrgElbay93	AJ130925	PcodAK93(AJ130926)	Snow et al., (1999)	WA, USA	47.622	-122.638
		Onchorhychus kisutch	*Makah	X59241	__	Bernard et al., (1992)	Makah National Fish Hatchery, Neah Bay, WA, USA	48.290	-124.650
		Paralichthys olivaceus	*JF00Ehi1	AB490792	__	Ito et al., (unpub.)	Ehime, JPN	33.750	132.600
		" "	*KRRV9822	AB179621	__	Byon et al., (unpub.)	" "	" "	" "
		" "	*KJ2008	JF792424	__	Kim and Kim (unpub.)	S. KOR	__	__
		Salmo salar	99-292	HQ453211	__	Thompson et al., (2011)	Discovery Is., BC, CAN	50.269	-125.160
	IVb	*Aplodinotus grunniens*	*U13653	HQ453209	DQ427105, ON55(HQ415763), Great Lakes(GQ255380)	Thompson et al., (2011)	Bay of Quinote, L. ON, CAN	44.150	-75.250
		Diporeia spp.	*HU54-M	HQ214135	MI27-M(HQ214134), MI18-M(HQ214133)	Faisal and Winters (2011)	L. Michigan, MI, USA	42.371	-87.113
		" "	*ON41	HQ415762	ON55(HQ415763)	" "	L. Ontario, CAN	43.717	-78.944
		Esox masquinongy	*MI03GL	GQ385941	*Myzobdella*(GQ255380)	Ammayappan et al., (2009), Faisal and Schulz (2009)	L. Michigan, MI, USA	42.371	-87.113
	IVc	*Fundulus heteroclitus*	*CA-NB00-01	EF079895	__	Gagné et al., (2007)	Ruisseau George Collette, near Bouctouche, NB, CAN	46.450	-64.682
		Salmo trutta	*CA-NS04-01	HQ168409	CA-NB00-02, CA-NB02-01, CA-NB0401b (HQ168406-HQ168408)	" "	French R., NS, CAN	46.485	- 60.441
Nv	**I**	*Oncorhynchus kisutch*	*Hededam	Z93412	__	Stone et al., (1997)	Spjarup Hededamme, DNK	56.648	9.271
	"I"	*O. mykiss*	*FiA03a_b.01	AM086360	FiA01a_b.00(AM086354), FiP02a_b.00(AM086356), FiP03.00(AM086357), FiA12.02(AM086373), FiA13.02(AM086374), FiA03.03(AM086379), FiA03.04(AM086382), FiA19.04(AM086383), , FiP04.01, FiA05.01, FiA04.01, FiP03.01,	Raja-Halli et al., (2006)	Åland Is., FIN	60.170	19.910

				FiA03.02, FiA04.02, FiA06.02, FiA07.02, FiA08.02, FiA09.02, FiA10.02(AM086361-AM086371), FiA14.02, FiA15.02, FiA16.02 (AM086375- AM086377)				
	" "	FiA02a.01, *FiA02.b01, FiA02ab01, FiA03ab01, FiA18.03, *FiA17.02, *FiA11.02, *FiP01ac00	AM086358-60, AM086380, AM086378, AM086372, AM086355	—	" "	" "	" "	" "
	" "	*FiU01.03	AM086381	—	" "	FIN	—	—
	" "	F1	U47848	—	Kurath et al., (1997)	" "	—	—
Ia	*Onchorhychus mykiss*	*Fil3	NC_000855	Fil3(X73873)	Schutze et al., (1999)	Baltic S., DEU	50.000	10.000
	" "	*French 0771	AJ233396	Strain0771(U28746)	Xing et al., (unpub.), Basurco and Benmansour (1995)	Seine-Maritime, FRA	49.500	0.500
	" "	*DK-6137, *DK-3592B	DQ159199, DQ159198	—	Einer-Jensen et al., (2004)	DNK	—	—
	" "	*AU-8/95	DQ159197	—	" "	AUS	—	—
	" "	*Strain 14-58	AF143863	—	Betts and Stone (2000)	FRA	—	—
	Salmo trutta	*FR-2375	FN665788	Strain2375(DQ159196)	Biacchesi et al., (2010), Einer-Jensen et al., (2005)	FRA	—	—
Ib	*Clupea pallasii*	*UKMLA986HE	GQ184827	DK-4p37(FJ460590), Strain 96-43(AF143862), SE-SVA-14(GQ184824)	Campbell et al., (2009), Betts and Stone (2000)	English Channel, GBR	50.600	0.000
	" "	*DK-1p8	GQ184826	—	Campbell et al., (2009)	DNK	—	—
	Gadus morhua	*Cod Ulcus	Z93414	—	Stone et al., (1997)	" "	—	—
	Gaidropsarus mediterraneus	*DK-1p40	DQ159200	—	Einer-Jensen et al., (2005)	" "	—	—
	Limanda limanda	*DK-5e59	GQ184825	—	Campbell et al., (2009)	Kattegat, DNK	56.320	11.560
	Onchorhychus mykiss	*SE-SVA-1033	FJ460591	—	" "	" "	" "	" "
Ie	" "	*GE-1.2	DQ159201	—	Einer-Jensen et al., (2005)	GEO	41.688	41.884
II	*Clupea pallasii*	*DK-1p49	DQ159193	—	" "	Baltic S., DEU	56.500	19.000
	Sprattus sprattus	*DK1p55	DQ162801	DK1p53(DQ159195), DK1p52(DQ159194)	" "	" "	" "	" "

III	*Melanogrammus aeglefinus*	*UK-H17/2/95	DQ159202	__	" "	North S., E. Shetland	60.283	0.044
	Onchorhychus mykiss	*KH230308-2	FJ362513	BV060408-52(FJ362510), FA281107(EU481506), V230308-5 (FJ362515)	Duesund et al., (2010)	Storfjorden Møre and Romsdal, NOR	62.360	6.160
	" "	H050308-2	FJ362512	__	" "	" "	" "	" "
	Scophthalmus maximus	*UK-860/94	DQ159203	__	Einer-Jensen et al., (2005)	Gigha, W. SCO	55.670	5.800
IVa	*Oncorhynchus kisutch*	*Makah	U28745	US Makah(DQ159204)	Basurco et al., (1995)	Makah National Fish Hatchery, Neah Bay, WA, USA	48.290	-124.650
	" "	*JF00Ehi1	AB490792	__	Ito et al., (unpub.)	Ehime, JPN	33.750	132.600
	Paralichthys olivaceus	*KRRV9822	AB179621	__	Byon et al., (unpub.)	JPN	__	__
	" "	*KJ2008	JF792424	__	Kim and Kim (unpub.)	S. KOR	__	__
IVb	*Esox masquinongy*	*MI03GL	GQ385941	__	Ammayappan and Vakharia (2009)	L. St. Clair, MI, USA	42.391	-82.911

Chapter 3

Gene diversification in an emerging quasispecies: A decade of mutation in fish Viral Hemorrhagic Septicemia (VHS) across the Laurentian Great Lakes

Evolutionary Biology. Stepien, C.A., Pierce, L.R., and Leaman, D.W.

3.1 ABSTRACT: Viral Hemorrhagic Septicemia virus (VHSv) is an RNA rhabdovirus that infects >80 marine and freshwater fish species across the Northern Hemisphere. A new and novel substrain (IVb) appeared in Lake St. Clair of the Laurentian Great Lakes in 2003, producing massive fish kills, spreading rapidly, and diversifying in a quasispecies mode over the next decade. We compare its evolutionary rates, mutation types, and biogeographic distribution for four genes – glycoprotein (*G*), non-virion (*Nv*), phosphoprotein (*P*), and matrix (*M*) – in relation to possible pattern and function. We analyze 12 VHSv-IVb isolates with unique *G*-gene haplotypes for mutations in the other genes. All 12 also are found to have unique *Nv*-gene sequences, evolving the fastest at $2.0x10^{-3}$ substitutions/site per year, with the most nucleotide substitutions (48); 60% are non-synonymous. The *G*-gene evolves at ~1/10 that rate ($2.1x10^{-4}$) and has the second most substitutions (14, 8 (57%) non-synonymous). The *P* and *M* genes evolve ~1.4 ($\mu=1.5x10^{-4}$) and 1.8x ($\mu=1.2x10^{-4}$) more slowly than *G*, with just 4 and 6 haplotypes. Haplotype changes among the four genes coincide in some isolates but seem independent

in others. Collective genetic diversity appears greatest in Lakes Erie and Ontario, close to the VHSv-IVb first identification site and nearer its hypothesized evolutionary origin. This rapid evolutionary diversification fits a quasispecies pattern for the *G* and *Nv* genes, which likely allows new viral variants to evade host recognition and immune responses.

3.2 INTRODUCTION

Due to small genomes, lack of polymerase proof-reading, and short replication times, RNA viruses have some of the most rapid known evolutionary diversification rates, reaching a nucleotide substitution per round of replication (Holmes, 2009). Some quickly diversify into swarms of closely related variants, termed quasispecies (Domingo et al., 2002; Lauring and Andino, 2010), as described for a new and novel substrain (IVb) of the fish Viral Hemorrhagic Septicemia virus (VHSv) that appeared a decade ago (in 2003; Elsayed et al., 2006) in the Laurentian Great Lakes (Pierce and Stepien, 2012). The resultant diversity of new variants may allow the viral population to evade host immune responses, improving its chances for persistence and spread to new areas and host species (Jaag and Nagy, 2010). VHSv is a single-stranded RNA rhabdovirus that kills >80 species (OIE, 2013) and IVb has spread rapidly to new locations in the Great Lakes and beyond. Since that time, fewer numbers of outbreaks have occurred, however a total of 17 new VHSv glycoprotein (*G*) gene variants were identified, including one by Groocock et al. (2007), 10 by Thompson et al. (2011), 3 by Cornwell et al. (2012), 2 by the USGS Western Fisheries Research Center (Molecular Epidemiology of Aquatic Pathogens for VHSv – MEAP VHSv; http://gis.nacse.org/vhsv/#), and one by Garver et al. (2013). This

suggests that the virus is rapidly evolving, yet may have become less virulent, which is the ability of the organism to cause disease, i.e., its degree of pathogenicity (Shapiro-Ilan et al., 2005). VHSv thus provides an interesting model to analyze comparative evolutionary patterns among genes in an emerging quasispecies.

Clinical signs of VHSv infection include internal and external hemorrhaging, pale gills, exophthalmia (bulging eyes), erratic swimming, and ascites (i.e., fluid in the peritoneal cavity) (Winton and Einer-Jensen, 2002; Kurath, 2012). Viral particles (virions) are transported via fish waste, reproductive fluids, skin secretions, boating, ballast water, and fishing tackle (Meyers and Winton, 1995; Kane-Sutton et al., 2010). Virions live up to 13 days in the water (Hawley and Garver, 2008), with optimal survival temperatures of 9-12^{o}C (Smail, 1999) and replication occurring at 2-20^{o}C (Wolf, 1988), and can travel up to 2 km (Meyers and Winton, 1995). They thus disperse rapidly and widely in aquatic systems.

The VHS rhabdovirus is bullet shaped, ~12,000 nucleotides in length, and has six open reading frames of 3'*N-P-M-G-Nv-L*'5 (Fig. 3-1). The novel non-virion (*Nv*) gene is a synapomorphy that distinguishes all members of the genus *Novirhabdovirus*, which includes three other fish viruses recognized by the International Committee on Taxonomy of Viruses (2013): Hirrame rhabdovirus (HIRRv), Infectious Hematopoietic Necrosis virus (IHNv), and Snakehead rhabdovirus (SHRv). Other putative Novirhabdoviruses include: Eel virus B12 (EEv-B12) and C26 (EEv-C26), and Rio Grande cichlid virus (RGRCv) (Galinier et al., 2012). The *Nv* protein suppresses apoptosis at the early stage of viral infection, providing the virus with time to replicate inside host cells (Ammayappan and Vakharia, 2011). It also is hypothesized to aid in viral replication and pathogenicity

(Ammayappan et al., 2010), defined as the ability of an organism to cause disease i.e., harm its host (Shapiro-Ilan et al., 2005). This may allow the virus to persist longer in the host, with increased shedding and spread. The VHSv *Nv*-gene appears to evolve rapidly at $\mu = 1.3 \times 10^{-3}$ (Pierce and Stepien, 2012).

The (*G*) and nucleoprotein (*N*)-genes evolve much more slowly, by ~1/5 and 1/3 that of *Nv*, respectively (Pierce and Stepien, 2012). The *G*-gene encodes for host cell attachment and endocytosis (Kurath, 2012). The *N*-gene associates with negative- and positive-sense RNAs, modulating the balance between genome transcription and replication (Dietzgen, 2012). Phosphoprotein (*P*) is responsible for viral replication in host cells and inhibits the host's activation of interferon (i.e., blocks antiviral responses); the matrix protein (*M*) inhibits promoter activity of interferon (Pore, 2012) and assists in the viral budding process (Biacchesi et al., 2002). RNA polymerase (*L*) guides viral genome transcription and replication (Kurath, 2012). Here we compare the mutation rates and substitution patterns of the VHSv *G*-, *Nv*-, *P*- and *M*- genes, while the *L*-gene remains unstudied.

VHSv first was identified in 1938 in European salmonid aquaculture (described as "*Nierenschwellung* disease"; Schäperclaus, 1938), and now is classified as comprising four strains (I-IV; Snow et al., 1999), whose evolutionary trajectory and biogeographic range were analyzed by Pierce and Stepien (2012). That study discerned that VHSv originated from a marine ancestor in the Atlantic Ocean ~700 years ago, diverging into two primary clades: strain IV in North America and strains I-III in Europe. In the latter clade, strain II forms the basal taxon to a clade containing strains I and III, which then are sister taxa.

Strain IV (now classified as IVa; Elsayed et al., 2006) was isolated in 1988 from salmon spawning along the Northwest Pacific coast of North America (Brunson et al., 1989; Hopper, 1989), and then was found in Japan in the mid-1990s (Takano et al., 2000). In 2000, another substrain (now designated as IVc per Pierce and Stepien, 2012) was discovered off the coast of New Brunswick, Canada, infecting estuarine mummichog (*Fundulus heteroclitus*) and three-spined stickleback (*Gasterosteus aculeatus*) (Gagné et al., 2007). In 2005, a new and especially virulent substrain (IVb) appeared in the Great Lakes, traced to a 2003 muskellunge (*Esox masquinongy*) from Lake St. Clair (Elsayed et al., 2006), causing massive fish kills of muskellunge, walleye (*Sander vitreus*), yellow perch (*Perca flavescens*), freshwater drum (*Aplodinotus grunniens*), and round goby (*Neogobius melanostomus*) (Thompson et al., 2011).

The first major VHS-IVb outbreaks occurred in 2005, concurrently at Lake St. Clair killing muskellunge (USDA-APHIS, 2006) and the Bay of Quinte, northern Lake Ontario in freshwater drum and round goby (Groocock et al., 2007). The largest outbreak to date spanned across the lower Great Lakes in 2006, including Lake St. Clair, the Detroit River, Lake Erie, the Niagara River, Lake Ontario, and the St. Lawrence River, (USDA-APHIS, 2006; Groocock et al., 2007; Whelan, 2009; FishLakeErie.com) infecting >10 species. These included: walleye, yellow perch, freshwater drum, round goby, smallmouth bass (*Micropterus dolomieu*), largemouth bass (*Micropterus salmoides*), rock bass (*Ambloplites rupestris*), northern pike (*Esox lucius*), bluegill (*Lepomis macrochirus*), pumpkinseed (*Lepomis gibbosus*), black crappie (*Pomoxis nigromaculatus*), gizzard shad (*Dorosoma cepedianum*), common carp (*Cyprinus carpio*), silver redhorse (*Moxostoma anisurum*), shorthead redhorse (*M.*

macrolepidotum), and lake whitefish (*Coregonus clupeaformis*). In 2007, VHSv-IVb outbreaks occurred in several scattered inland lakes (Whelan, 2009), including: Little Lake Butte des Morts (WI) west of Lake Michigan, Lake Winnebago (WI) west of central Lake Michigan, Budd Lake (MI) – between Lakes Michigan and Huron, and Skaneateles Lake (NY) in the Finger Lakes south of central Lake Ontario (Fig. 3-2a). In 2008, VHSv killed round goby and yellow perch in Milwaukee Harbor, western Lake Michigan (Wisconsin Department of Natural Resources, 2008). Lake St. Clair had another VHS-IVb outbreak in 2009 in smallmouth bass (Faisal et al., 2012). Lake Superior has been outbreak-free, however, VHSv was detected in 2009 from a lake herring (*Coregonus artedi*) in the Apostle Islands at the location on Fig. 3-2a (Thompson et al., 2011). No VHSv die-offs were reported in 2010. In 2011, outbreaks occurred again in Milwaukee Harbor, Lake Michigan (yellow perch and yearling gizzard shad) and Budd Lake (smallmouth bass and sunfish) (Great Lakes Fish Health Committee, 2011). No subsequent outbreaks appear to have been reported. In 2012, two fish (a freshwater drum and largemouth bass *Micropterus salmoides*) from central Lake Erie (Sandusky OH) tested positive for our laboratory (Pierce et al., 2013; Pierce and Willey et al., 2013).

Pierce and Stepien (2012) hypothesized a quasispecies model of evolution for IVb, which has rapidly diversified from its original sequence (MI03GL) – found in muskellunge from Lake St. Clair in 2003 – into a "cloud" of closely related *G*-gene variants. Here we evaluate the diversification pattern of IVb over temporal and spatial scales, comparing *G*-gene variants with those in the *Nv, P,* and *M* genes.

The objective of the present study is to discern the evolutionary and biogeographic diversification of VHSv-IVb in the Great Lakes. We compare and contrast

the mutation patterns and rates of four genes – *G*, *Nv*, *P*, and *M* – asking: (1) How much genetic change has occurred?, (2) How are the mutations phylogenetically and geographically distributed?, and (3) Are these relationships consistent among the genes? Understanding how VHSv changes over time and space may aid prediction of its future spread and virulence.

3.3 MATERIALS AND METHODS

3.3.1 Samples, gene amplification, and sequencing

This study analyzes 12 VHSv-IVb isolates (labeled a-l in Table 3.1, Fig. 3-2) that have unique *G*-gene sequences; 10 of which were described by Thompson et al. (2011); (provided to us by G. Kurath, USGS, as purified RNA) and 2 here are identified here from fish in central Lake Erie (a largemouth bass and freshwater drum from Sandusky OH). The 10 *G*-gene haplotypes (labeled a-j; Table 3.1) were obtained from 108 fishes and invertebrates sampled in Lake Superior (1), Lake Michigan (7), Lake Huron (5), Lake St. Clair (14), Lake Erie (25), Lake Ontario (12), St. Lawrence Seaway (24), and inland waterbodies (20; Wisconsin, Michigan, Ohio, New York, and Ontario) (Thompson et al., 2011). Unique haplotypes k and l from two fish samples are sequenced by us here; these were swimming erratically and were collected by the Ohio Division of Wildlife Sandusky Fish Research Unit off their boat dock in April and May 2012.

Fish were euthanized with overdose of 25 mg/ml tricaine methanesulfonate (MS-222; Argent Chemical Laboratories, Redmond, WA) and decapitated, following our

University of Toledo Institutional Animal Care and Use Committee (IACUC) approved protocol #106419. The surgical site (anus to operculum) was disinfected with 100% ethanol and betadine using sterile equipment. The spleen was removed, placed in a 1.5 mL tube of RNAlater (Qiagen, Valencia, CA), and stored at -80^{o}C. Gloves and instruments were changed between fish to ensure sterility. Specimens were disposed of following the approved University of Toledo biohazard protocol.

Spleen tissues were ground under liquid nitrogen with a sterile mortar and pestle. RNA was extracted with the TriREAGENT® (Molecular Research Center Inc., Cincinnati, OH) manufacturer's protocol. RNA from each sample was re-suspended in 30 µl RNase-free water, quantified with a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA), and adjusted to 1 µg RNA/µl. RNA then was reverse transcribed to cDNA using 1 µg RNA, 5X First Strand buffer, 10 mM dNTPs, 0.05 mM random hexamers, 25 U/µl RNasin, and 200 U/µl M-MLV in a 90 µl reaction (rxn). Reverse transcription was carried out at 94°C for 5 min, 37°C for 1 h, and 94°C for 5 min; cDNA was stored at -20°C.

Each gene was amplified using 2 µl cDNA (30-60 ng DNA template) in a polymerase chain reaction (PCR) containing 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris-HCl, 50 µM of each dNTP, 0.5 µM each of the forward and reverse primers (Table 3.2), and 1 unit *Taq* polymerase (Promega, Madison, WI) in a 25 µl rxn. Reactions included initial denaturation of 2 min at 95^{o}C, followed by 30 cycles of 30 sec at 95^{o}C, 30 sec at primer and strain-specific annealing temperatures (Table 3.2), and 1.5 min at 72^{o}C, with a final extension of 2 min at 72^{o}C.

Amplification products from PCR were visualized on 1% agarose mini-gels stained with ethidium bromide. Successful rxns were gel purified using QIAGEN Purification Kits. Samples having less efficient amplification were PCR-purified. Purified products were assessed on mini-gels and DNA sequencing was outsourced to Eurofins MWG Operon (http://www.operon.com/default.aspx). All sequences were checked, identified, and aligned by us with BIOEDIT v7.05 (Hall, 1999), and deposited in GenBank (http://www.ncbi.nlm.nih.gov/genbank/ as Accession numbers XXXXX-XXXXX). Our final aligned sequences encompassed: 669 basepairs (bp) for the *G*-gene, 366 for the *Nv*-gene, 541 for the *P*-gene, and 564 for the *M*-gene.

We additionally analyzed all available VHSv-IV sequences for the *Nv*, *P*, and *M* genes from GenBank and the literature, and a subset of the 213 IVa and 5 IVc sequences for the *G*-gene (the entire set was previously reported by Pierce and Stepien, 2012). Homologous sequences were recorded, pruned, and geographically referenced (Appendix 3.1). The original MI03GL isolate was used as the reference sequence (labeled "a"; Table 2). Unique haplotypes, collection year(s), and watershed(s) were identified and used for our phylogenetic analyses, mutation comparisons, and haplotype networks. Representative sequences from VHSv strains I (GenBank Accession # Z93412), II (# HQ112247, DQ159193, XXXXXX), and III (# FN665788) were used as outgroup comparisons and to root the phylogenetic trees.

3.3.2 Genetic diversity, substitutions, and mutation rates

Haplotypic diversities (H_D) for the combined four gene regions and each individual gene were determined with ARLEQUIN v3.5.1.3 (Excoffier and Lischer, 2010). Numbers of transitional and transversional substitutions, and uncorrected pairwise (*p*-) distances, among the haplotypes were determined in MEGA v5.0 (Tamura et al., 2011), and numbers of synonymous and non-synonymous changes per nucleotide site by DNASP v5.10.01 (Librado and Rozas, 2009). Substitution types were plotted against the *p*-distances using EXCEL (Microsoft Corp.), correspondence to linear models was evaluated with *F*-tests (Sokal and Rohlf, 1995), and ANCOVA in R v2.15.2 (R Development Core Team, 2012) was used to compare the relationships. Student's *T*-tests were employed to evaluate the relative numbers of substitution types. Tajima's (1989) *D* in ARLEQUIN tested for possible selection.

Mutation rates, measured as the number of nucleotide substitutions per site per year (μ=substitutions site^{-1} yr^{-1}), were calculated using divergence estimates from BEAST v1.71 (described below; Drummond et al., 2012) and the *p*-distances. Evolutionary rates were estimated for each gene region and the overall combined gene data set.

3.3.3 Phylogenetic relationships and divergence patterns

Patterns of genetic aggregation, spatial dispersion, and number of evolutionary steps were evaluated with statistical parsimony haplotype networks in TCS v1.21

(Clement et al., 2000). For those requiring additional analyses to discern relationships among haplotypes, connection limits with a maximum of 500 steps were assigned. We tested the hypothesis of genetic isolation ($\theta_{ST}/1-\theta_{ST}$) by geographic distance (the natural logarithm of nearest waterway geographic distance (km)) using Mantel tests in GENEPOP v4.2 (Rousset, 2008). The nearest roadway distance was used for inland-locked Budd Lake (Clare County, Michigan). Regression significance was interpreted from 10,000 permutations.

Phylogenetic relationships were evaluated for each of the four genes using two approaches: maximum likelihood (ML) in PHYML v3.0 (Guindon and Gascuel, 2003) and Bayesian analyses in MRBAYES v3.2.1 (Ronquist and Huelsenbeck, 2003). The corrected Akaike Information criterion (AICc) from jMODELTEST v0.1.1 (Posada, 2008) was employed to determine the most appropriate nucleotide substitution model. jMODELTEST selected the TPM1uf (Posada, 2008) with a gamma distribution (α=0.2340) for the *G*-gene, TIM2 (Posada, 2008) plus invariant sites (I=0.2670) for the *Nv*-gene, TIM1 (Posada, 2008) with a gamma (α) distribution (α=0.3360) for the *M*-gene, and TVM (Posada, 2008) plus invariant sites (I=0.3700) for the *P*-gene.

ML nodal support was compared using 2,000 nonparametric bootstrap pseudoreplications (Felsenstein, 1985). Bayesian phylogenetic analyses were conducted in MRBAYES with a Metropolis-coupled Markov chain Monte Carlo (MCMCMC) approach and run for 5,000,000 generations, with sampling every 100 generations. The burn-in period was determined by plotting log likelihood values to identify when stationarity occurred; the first 25% (1,250,000 generations) were discarded as burn-in.

To discern the overall consensus phylogeny, we tested combinability of the four genes into a single data set using an Incongruence Length Difference test (ILD; Farris et al., 1995) in PAUP* v4.0 (Swofford, 2003). This included strain IV isolates that were sequenced for all four-gene regions, totaling 16 individuals (IVa=3, IVb=12, IVc=1) and 2,140 aligned bp, and single representatives from strain I (Hededam, GenBank Accession # Z93412) and strain III (strain 23-75, FN665788) were used to root the tree. ILD tests were based on 1,000 replicates in a parsimony framework, with a critical α=0.01. Results revealed that the four genes were combinable in a total evidence analysis (p>0.05). jMODELTEST selected the GTR model (Posada, 2008), with a gamma distribution (α=0.3110). ML and Bayesian analyses were performed, as above.

Comparative divergence times were estimated in BEAST using the same substitution model, gamma distributions for the combined, *G*-, and *M*-gene data sets, invariant sites for the *Nv*- and *P*-genes, a relaxed molecular clock that assumed a lognormal distribution, and 10,000,000 generations, with parameters sampled each 100 generations. Outputs were assessed in TRACER v1.5 (in BEAST) to ensure that values reached stationarity; those that did not were re-run using 50,000,000 generations. Collection/ isolation dates were used as calibration points for each sample.

3.4 RESULTS

3.4.1 Genetic diversity and mutations in the G-, Nv-, P-, and M-genes

The 12 VHSv-IVb isolates are lettered a-l in Table 3.1; all have unique sequences for the combined genes, *G*-, and *Nv*-gene datasets, designated *C*-a through *C*-l, *G*-a to *G*-l, and *Nv*-a through *Nv*-l. The *P*-gene has just 4 haplotypes (*P*-a, *P*-b, *P*-d, *P*-k) and the *M*-gene has 6 (*M*-a, *M*-b, *M*-h, *M*-j, *M*-k, *M*-l). These have just 50% and 30% of the diversity of the *G* and *Nv* haplotypes.

All substituted nucleotide positions are substituted only once in all genes, indicating lack of saturation (i.e., no multiple "hits") (Table 3.3). The *Nv*-gene has the most substitutions (48/366 positions; 13.1%), followed by *G* (14/669; 2.1%), *P* (6/539; 1.1%), then *M* (5/564; 0.9%) (Table 3.4). Most *Nv*-gene substitutions are at third codon "wobble" positions (21/48, 44%), which greatly outnumber those at the first (14/48, 29%) and second (13/48, 27%). The *G*-gene has more nucleotide changes at the second codon position (7/14, 50%) than the first (3/14, 21%) and third (4/14, 29%). Most substitutions in *P* and *M* are at the third codon position (4/6, 67%, 4/5, 80%, respectively). *P* has no changes at the second position, and *M* has none at the first (Tables 3.3 and 3.4).

Nucleotide substitutions increase linearly with *p*-distance (likely positively associated with time) in the combined data set (R^2=0.97-1.00, $F_{1,64}$=2,313-1.46x10^8, p<0.001), and the *Nv* (R^2=0.89-0.90, $F_{1,64}$=530.7-577.0, p<0.001) and *P* genes (R^2=0.65-0.77, $F_{1\text{-}4}$=7.5-13.0, p=0.02-0.05) (Fig. 3-3a,c,e). Substitutions appear level across the *G* and *M* genes (Fig. 3-3b,d). Transitional substitutions significantly outnumber transversions in the combined data set (t=10.65, df=130, p<0.001), and the *G* (t=5.96, df=130, <0.001), *Nv* (t=9.18, df=130, p<0.001), and *M* (t=4.81, df=28, p<0.001) genes (Fig. 3-3) – but not for *P*. Patterns of transitions and transversions do not appear directly related in any of the gene data sets (ANCOVA: $F_{8\text{-}128}$=8.58-1,717, p<0.001-0.01).

Numbers of synonymous and non-synonymous substitutions do not significantly differ, except in the *Nv*-gene (t=2.02, df=130, p=0.045) (Fig. 3-4). Synonymous and non-synonymous changes show different trajectories, in relation to increasing genetic distances (ANCOVA: $F_{8\text{-}128}$=2.8-333.2, p<0.001-0.05). The *Nv*-gene has the most synonymous (19) and non-synonymous (29) changes, followed by *G* (6, 8), *P* (3, 3), and *M* (2, 3) (Table 3.4). Nucleotide substitutions follow a linear trend with *p*-distance for the combined data set (R^2=0.79-1.00, $F_{1,64}$ =242.1-1.10x10^8, p<0.001) and the *Nv* (R^2=0.86-0.93, $F_{1,64}$ =382.6-843.6, p<0.001) and *P* genes (R^2=0.79-0.86, $F_{1,4}$=15.4-24.0, p=0.01-0.02) (Fig. 3-4a,c,e). The *G*- and *M*-gene relationships do not fit a linear prediction (Fig. 3-4b,d). Tajima's *D* values are negative and significant for all (combined data set: *D*=1.76, p=0.03; *G*-gene: *D*=-2.07, p=0.003; *Nv*-gene: *D*=-1.58, p=0.05; *P*-gene: *D*=-1.89, p=0.01) except *M*, indicating possible purifying selection. The *G*-gene has the strongest indication of selection.

Overall evolutionary rate of VHSv-IVb is estimated at μ=4.7x10^{-4} for the combined data set. Individual gene rates are: μ=2.1x10^{-4} for *G*, μ=2.0x10^{-3} for *Nv*, and μ=1.5x10^{-4} for *P*. An unresolved relationship in our *M*-gene phylogenetic tree circumvents calculation of its evolutionary rate.

3.4.2 Genetic divergences and relationships among haplotypes

The 12 *G*-gene haplotypes differ by 1-7 bp, averaging 2.8±0.04 pairwise nucleotide substitutions (mean uncorrected *p*-distance=0.004±0.002; range=0.001-0.010). Mean difference between *Nv*-gene haplotypes is ~3.6x greater, differing by a mean of

10.7±2.3 bp, with sequences having 1-31 substitutions (*p*-distance=0.028±0.008, range=0.003-0.085). *P*-gene haplotypes diverge by a mean of 2.5±0.7 bp, ranging from 1-4 differences among sequences (*p*-distance=0.006±0.003; range=0.002-0.009). The *M*-gene has the fewest number of substitutions, with a mean of 1.9±0.3 bp (1-3 bp) (*p*-distance=0.003±0.002; range=0.002-0.005) (Table 3.4).

Just three isolates possess unique sequences for all four genes: a (MI03GL), b (vc002), and k (LMB) (Table 3.2). Of these, haplotye *G*-a is known to be the most widespread, found in all five Great Lakes (Fig. 3-2a; Thompson et al., 2011). *G*-b is the second most widespread in Lakes Michigan, Huron, Erie, and Ontario. *G*-k is known only from a single fish in central Lake Erie, uniquely identified here. We also describe a unique *G*-sequence for another central Lake Erie fish from the same location (*G*-l), which has unique *Nv* and *M*-haplotypes, but whose *P*-gene is identical to isolate a (Table 3.1).

In the combined gene haplotype network (Fig. 3-5a), as well as the *Nv*-gene network (Fig. 3-5c), haplotypes a, k, and l are located away from the others, with l being the most divergent. Six steps from a, nine haplotypes stem in a star-like pattern (b-j), with c being the most different. The *G*-gene network (Fig. 3-5b) depicts abundant haplotype a as central, from which eight differ by a single step. A second abundant haplotype, b, differs by a single nucleotide from a, and serves as another center. Haplotypes o, p, q, and s (from Lake Ontario in 2009-2010) form a cluster one step away from b. The *P* and *M* genes also depict abundant type a in the center of the networks (Fig. 3-5d,e). All other *P* and *M* haplotypes stem from a by 1-3 steps, with *P* averaging 2.5±0.7 steps and *M* averaging 1.9±0.3 (Table 3.4). Pairwise genetic distances between the combined

haplotypes do not correspond to genetic isolation by geographic distance pattern, according to Mantel tests.

The combined gene phylogeny resolves two primary IVb clades (marked 1 and 2 on Fig. 3-6a), estimated to have diverged ~15 years ago (ya) (9.1-23.2 highest posterior density; HPD). Clade 1 diversified ~9 ya (9.0-11.7 HPD), with *C*-a as the sister taxon to a clade of *C*-k and *C*-l (labeled *C*-1.1). The latter differ from each other by 6 nucleotides, diverging ~3 ya (0.002-6.9 HPD). Clade 2 originated ~14 ya (7.2-23.0 HPD) and contains the most haplotypes. Within it, *C*-h and *C*-d are a sister group (*C*-2.1) that diverged ~7 ya (6.0-10.6 HPD) by 9 nucleotides; they are geographically distant, with *C*-h in Lake Superior and *C*-d in eastern Lake Erie. Located in clade 2 and from eastern Lake Erie and Lake Ontario, *C*-c is the most divergent haplotype by >23 bp. All other combined haplotypes differ by 8-20 changes from *C*-a.

G-gene tree (Fig. 3-6b) relationships show a basal unresolved trichotomy for IVb, comprising haplotypes i, b, and clade 1 (a, c-h, j-l). Within clade 1, k and l are sister taxa (1.1; these are 2012 samples from Sandusky OH, central Lake Erie) that differ by 2 nucleotides. The *Nv*-gene tree (Fig. 3-6c) shows two primary clades (1 and 2). Within clade 1, k and l are sister taxa (1.1), which again differ by 2 bp and comprise the sister group to a. Clade 2 contains haplotypes c-j in two internal subclades (2.1 and 2.2). Its subclade 2.1 has d and h as sister taxa, differing by 4 nucleotides and found one year apart in Lakes Erie (2006) and Michigan (2007), respectively (Fig. 3-2b). Subclade 2.2 groups c and b, differing by 26 bp and identified just one year apart in Lakes Erie (2007) and Michigan (2008). Haplotype c is the most divergent in the entire IVb *Nv*-gene phylogeny.

No major groupings are evident for the *P*-gene (Fig. 3-6d), which has just 5 haplotypes. Haplotype b is the most divergent, by 4 bp. A single clade (1.1) is supported in the *M*-gene tree (Fig. 3-6e), which again links haplotypes k and l, differing by one substitution. This same sister grouping thus characterizes the combined, *G*, *Nv*, and *M* trees (Fig. 3-6a,b,c,e). The three other *M*-haplotypes differ by just a single bp from haplotype a on the tree (Fig. 3-6e); these were found in Lake Michigan in 2007-2008.

3.5 DISCUSSION

3.5.1 Comparative mutation rates and diversity in the four genes (Question 1)

VHSv-IVb has diversified extensively during its decade in the Great Lakes, with the *Nv*-gene evolving the fastest at $\mu = 2.0 \times 10^{-3}$, ~10 times faster than the *G*-gene, ~14x than *P*, and ~17x than *M* (1.2×10^{-4}, estimated from Fig. 3-3 and Fig. 3-4). Our previous work estimated that the *N*-gene evolves at 4.3×10^{-4} across all four VHSv strains (Pierce and Stepien, 2012), which thus is 3x slower than the *Nv*-gene. The *N*-gene was not investigated here, but is in progress by Leaman and Stepien. Most VHS IVb *Nv*- and *G*-gene mutations are non-synonymous (60 and 57%), which may result in functional protein changes that are acted upon by selection. Our Tajima's (1989) *D* tests indicate that the *G*, *Nv*, and *P* genes may be under purifying (=negative) selection, i.e., the selective removal of deleterious alleles to maintain the long-term stability of their protein structure. Central *G*-gene sequences for the related *Novirhabdovirus* IHNv also appeared to be subject to purifying selection (Troyer and Kurath, 2003), but later analyses of the

entire gene indicated possible positive selection (Padhi and Verghese, 2008). Wu et al. (2010) stated that functional constraints in genomes play a key role in their evolutionary rates, with their relative amounts of genetic variation influenced by protein length, codon usage, gene expression, and protein interaction. Such differences among VHSv genes likely regulate their mutation rate and function.

Reverse genetic (knockout gene) studies of the VHSv *Nv*-gene suggest it plays an anti-apoptotic role in early infection (Ammayappan and Vakharia, 2011). This may lead to longer persistence of the virus in the host, corresponding to higher probability for new mutations to arise. Cell culture experiments by Biacchesi et al. (2010) discerned that VHSv replication was reduced 10,000-fold when the *Nv*-gene was deleted. Ammayappan et al. (2010) hypothesized that *Nv* functions in viral replication efficiency and pathogenicity, since VHS-IVb challenged yellow perch experienced increased mortality compared to an altered strain in which *Nv* was replaced with a green fluorescent protein. Similar findings of *Nv* involvement in viral replication and pathogenicity were reported for IHNv (Biacchesi et al., 2000; Thoulouze et al., 2004). In contrast, SHRv *Nv* deleted mutants had analogous replication and pathogenicity to those of the wild type (Johnson et al., 2000; Alonso et al., 2004). This may represent an evolutionary reversal, in which function may have reverted to an ancestral "neutral" condition for *Nv* in Novirabdoviruses. Among the Novirabdoviruses, SHRv is the more closely related to VHSv and IHNv is more ancestral (Ammayappan and Vakharia, 2009; Pierce and Stepien, 2012).

The *G*-gene evolves at ~2.1×10^{-4}, ~1/10 the rate of *Nv*. Holmes et al. (2002) estimated a similar rate of 4.1×10^{-4} for rabies, which is a rhabdovirus distantly related to

the Novirhabdoviruses. The *G*-gene encodes the glycoprotein that forms spike-like projections on the viral particle surface (Fig. 3-1), which attaches to the host cells by binding to the cellular receptor fibronectin (Bearzotti et al., 1999). The glycoprotein also has an antigenic role, eliciting strong innate host immune responses (Kurath, 2012). Mutations in the *G*-gene may facilitate adherence and entry into the host's cells in the face of the host's ability to recognize the virus and circumvent that entry.

Numbers of transitional substitutions greatly outnumber transversional for all genes except *P*, indicating that their phylogenetic signal is free of saturation (see Stepien and Kocher, 1997; Hicks and Duffy, 2011). The *P*-gene functions in VHSv viral replication (Ammayappan and Vakharia, 2011) and has been implicated to inhibit the host's innate immune response (Pore, 2012). We find that *P* has low sequence divergence and evolves much slower than *Nv* and *G*. Changes in the *P*-gene may be deleterious to viral replication.

The *M*-gene evolves the slowest, and appears to be involved in inhibiting protein expression of the infected host cells (Pore, 2012). *M*-gene proteins also interact with those from the *G*-gene to assist in viral budding (Biacchesi et al., 2002). In another rhabdovirus, Vesicular Stomatitis virus (VSv), the *M* protein inhibits nuclear transport of mRNA and blocks translation of host proteins (Fontoura et al., 2005), allowing host cellular machinery to replicate more virus. This interaction may also be occurring in VHSv-IVb. An alternative hypothesis is it may play a role in last step of the lytic cycle by facilitating release of viral particles from host cells. Mutations in the *M*-gene might lead to less efficient signal transduction or interfere with viral budding. Reverse genetics research is needed to test these hypotheses.

3.5.2 Geographic patterns and phylogenetic relationships from G-gene sequences (Question 2)

Lake Erie houses the greatest known VHSv-IVb *G*-gene diversity (8 haplotypes, 5 of which appear unique; 63%), followed by Lake Ontario (7, 3; 43%), Lake Michigan (6, 4; 67%), Lake Huron (3, 1; 33%), Lake St. Clair (1,0), and Lake Superior (1,0), based on our sampling and that by Garver et al. (2013), and MEAP VHSv (2013; http://gis.nacse.org/vhsv/#). Lake Erie is the warmest and shallowest of the Great Lakes (Makarewicz and Bertram, 1991), with its warmest and shallowest western basin located relatively close to the original VHSv-IVb isolate "a" (MI03GL) from Lake St. Clair. Viral diversity likely has been facilitated by Lake Erie's warmer water temperatures and high fish abundance, estimated to comprise 80% of all Great Lakes fish numbers (J. Reutter, pers. commun.). Kane-Sutton et al. (2010) analyzed yellow perch from central Lake Erie in 2007 and 2008, discerning that VHSv-IVb infection was greatest at water temperatures of 12-18^{o}C and with higher fish densities, such as during spawning aggregations. The 2006 VHSv-IVb outbreak occurred in June at surface water temperatures of ~15-16^{o}C, when many fishes, including yellow perch, were spawning.

Haplotypes a and b are the oldest (2003, 2005 respectively), most abundant (61, 36 sequenced occurrences), and widespread; a has been identified in all of the Great Lakes and b from Lakes Michigan, Huron, Erie, and Ontario (Fig. 3-2a). Haplotype b was first identified from the 2005 Bay of Quinte, Lake Ontario outbreak. The two haplotypes do not appear closely related in *Nv*- and *P*-gene sequences, and differ by a single step in the *G*- and *M*-gene haplotype networks. The two comprise the center nodes in the *G*-gene

network, with many other haplotypes diverging from them. Haplotype b may have differentiated from a, subsequent to the first discovery of VHS IVb in the Great Lakes. Alternatively, according to the *Nv*-gene network, both a and b may have had separate origins dating to several years prior to the outbreak; these may have descended from a common ancestor in the western North Atlantic Ocean that invaded the Great Lakes (see Pierce and Stepien, 2012).

Eleven *G*-gene haplotypes (b-h, j-m) appear closely related to the original a haplotype (Fig. 3-5b). Two descendants of haplotype a (d in western Lake Erie OH and m in central Lake Erie) appear to date to the 2006 outbreak. Five haplotypes date to 2007 (e and f from the outbreak in Budd Lake MI, h in northwest Lake Michigan, g in western Lake Erie, and c in eastern Lake Erie NY and south of Lake Ontario in the inland NY Finger Lakes). Haplotype j was identified from the 2008 outbreak near Milwaukee WI in Lake Michigan.

Seven haplotypes (i, n-s) appear to have descended from haplotype b, which were found from sampling in 2008 (i, in the Milwaukee WI Lake Michigan outbreak), 2009 (o, p, and q in eastern Lake Ontario), and 2010 (s from Lake Simcoe, ON, and n in northwest Lake Huron). Occurrence of new *G*-gene variants stemming from a and b suggests that the virus likely was evolving after the 2006 outbreak throughout those Great Lakes; this evolution may have been in response to the host populations developing resistance to VHSv following the large fish die-offs. We here describe new k and l haplotypes sampled from central lake Erie (Sandusky OH) in 2012. It is likely that haplotype l descended from k, as shown in the *G* network (as well as the *Nv*- and *M*-gene networks; Fig. 3-5c,e). Continued evolution of new *G*-gene variants may facilitate the virus' ability to avoid

recognition by new host individuals despite previous exposure of the host population to other viral variants.

Other VHSv studies that analyzed the relationships among *G*-gene IVb haplotypes reported similar findings, but lacked our k and l samples and did not analyze other genes. Specifically, analyses by Thompson et al. (2011) and Cornwell et al. (2012) found that haplotypes a, c-h, and j group together in an internal node, with haplotypes b and i located on the periphery. Those findings appear analogous to the unresolved basal trichotomy seen in our *G*-gene tree. Our study employs a more rigorous phylogenetic approach and analyzes other genes, providing a more robust analysis of VHSv-IVb evolutionary patterns.

3.5.3 Evolutionary patterns: consistency and differences among the genes (Question 3)

In the *Nv*-gene network (Fig. 3-5c), haplotypes a, k, and l are distinct from the others. The 2012 k and l isolates are very different from all but a. These mutations were not linked to an outbreak, but continued evolution of *Nv*-gene variants, which may be selected to persist longer in the host and facilitate viral replication. The k and l variants also have unique *M*-gene sequences (two of six haplotypes) and k also varies in its *P*-gene sequence (one of just 5 haplotypes). These mutations in such highly-conserved genes suggest that VHSv-IVb is evolving to aid its replication and avoid the innate immune response of the host.

The *Nv* haplotype network shows two other haplotypes, b and c, which also are very different from the others (Fig. 3-5c). Haplotype b is distinct in the *G*, *Nv*, *P*, and *M*

genes; it is at the center of the *G*-gene network and appears to be one of the oldest and original VHSv-IVb variants. Haploype c is not different than a in the *P* and *M* genes; it is from eastern Lake Erie and the New York Finger Lakes south of Lake Ontario.

In the *Nv* network, a distinct and closely related cluster of seven haplotypes (d-j) diversified in 2006-2008 in Lakes Erie and Michigan. This cluster contains two Lake Erie haplotypes (d and g), respectively from the 2006 outbreak and 2007. *Nv* cluster members appear to have been very successful in the 2007 Budd Lake outbreak and the 2008 Milwaukee Lake Michigan outbreak. In the *G*-gene network, six of these *Nv*-gene cluster members (all except for i) descend from the central *G*-a haplotype. All of these *Nv* cluster members thus appear to be evolutionarily related, stemming from a common *Nv* mutation six nucleotide steps from isolate a. They then rapidly continued to mutate in a star-like cluster, with 1-4 additional substitutions. A single member of the cluster – d – also mutated in the *P*-gene (western Lake Erie 2006) and two others had *M*-gene mutations – h and j (Lake Michigan, 2007 and 2008). All of the other members of the *Nv*-gene cluster members have the ancestral "a" sequences.

3.5.4 Gene diversification and evolution of VHSv-IVb as an emerging quasispecies

It is interesting to note that haplotypes k and l recovered in this study were collected from the same location (Sandusky OH, central Lake Erie), just one month apart in 2012, yet have unique sequences for the *G*, as well as the *Nv* and *M* genes. Those fish had no outward lesions or internal hemorrhaging, with fewer clinical signs of infection than those that characterized the 2006 fish kill. Their symptoms were disorientation and

erratic swimming. Other fish in the area did not appear affected. This suggests that Lake Erie fish had developed resistance to the IVb strain, while some mutated variants had moderate success.

Monitoring (Kane-Sutton et al., 2010; Frattini et al., 2011) and sequencing studies (Thompson et al., 2011; Cornwell et al., 2012; Faisal et al., 2012) have reported that some Great Lakes fish that tested positive for VHSv lacked clinical signs of the disease. For example, Frattini et al. (2011) surveyed New York State fish (in eastern Lake Erie, Lake Ontario, and inland lakes) for VHSv in fall-winter 2006, following the summer 2006 outbreak. They found that 69 of 1,011 (6.8%) samples tested positive, but lacked visual VHSv symptoms. Tests of Lake Erie yellow perch in 2007-2008 (Kane-Sutton et al., 2010) showed that 60% of pooled samples tested VHSv positive, yet appeared healthy and disease free. Thompson et al. (2011) analyzed 108 fishes and invertebrates (*Myzobdella lugubris l*eech and amphipod *Diporeia* spp.) sampled in 2003-2009 across the Great Lakes, finding that 28 (26%) of the individuals were positive, but asymptomatic. This supports the hypothesis that VHS-IVb is prevalent across the Great Lakes, but has become less virulent over time. Our present results suggest that the virus has responded by mutating into a variety of variants.

Since VHSv-IVb was first discovered in 2003, it has evolved into at least 19 unique *G*-gene haplotypes, with Lake Erie housing the greatest known diversity. The *G*-gene encodes attachment to host cells and induction of endocytosis. We show that the *Nv*-gene evolves ~10x faster; its rapid mutation rate may suppress apoptosis in early viral infection, and play a role in pathogenicity. The *P*-gene evolves more slowly, ~87% slower than the *Nv*-gene and ~43% less than *G*; it is believed to function in viral

replication and to suppress interferon induction in host cells. The *M*-gene evolves ~17x more slowly than the *Nv*-gene, ~1.8x than the *G*-gene, and 1.3x than the *P*-gene. Its evolutionary change appears constrained by its role in budding new viral particles; mutations could reduce viral production.

Of the 12 isolates we tested, haplotypes a and b are centrally located in the *G*-gene network, with other haplotypes stemming from them. These two haplotypes also have closely related *M*-gene sequences. However, in the *Nv* and *P* networks, a and b appear more distant. The k and l haplotypes (2012) in central Lake Erie are evolutionary divergent from the other haplotypes in three of the genes (*G*, *Nv*, and *M*), and are closely related to each other in all gene sequences. There is a cluster of seven *Nv*-gene haplotypes identified in 2006-2008 from Lakes Erie and Michigan; these radiate outward from the original haplotype by seven steps. In conclusion, haplotype mutational patterns appear to coincide among genes in some cases, but appear independent for others.

This is the first study to sequence the *Nv*, *P*, and *M* genes of VHSv-IVb, beyond that of the original MI03GL (a) Great Lakes isolate. We compare and contrast the evolutionary rates and mutational patterns of IVb, describing rapid diversification in the *G* and *Nv* genes, most of which are non-synonymous mutations. These results increase fundamental understanding to begin elucidating the co-evolutionary patterns of VHSv and its hosts, and may assist prediction of future spread and success.

3.6 ACKNOWLEDGEMENTS

Research funding included: NSF DDIG #1110495, USDA-NIFA (CSREES) #2008-38927-19156, #2009-38927-20043, and #2010-38927-21048, USDA-ARS #58-3655-9-748 A01, NOAA Ohio Sea Grant #R/LR-015, and NSF GK-12 #DGE-0742395. We thank present and past members of the Great Lakes Genetics/Genomics Laboratory for assistance, including: A. Haponski, S. Karsiotis, M. Neilson, V. Palsule, C. Prichard, and M. Snyder. G. Kurath of USGS (Seattle, WA), K. Garver of Canada Fisheries and Oceans (Ottawa, Canada), and T. Gadd of the Finnish Food Safety Authority Evira (Helsinki, Finland) provided samples. Sequence data were generated by D.L. laboratory members Jacob Blandford, Adam Pore, Julia Wildschutte, Tyler Williams, and C.S. laboratory member Vrushalee Palsule. P. Uzmann, M. Gray, and R. Lohner gave logistic support for the project. We thank Jennifer Sieracki for information on the occurrence and spread patterns of VHSv-IVb.

Table 3.1

Homologous VHSv-IVb haplotypes for the *P*- and *M*-genes, in reference to unique types for the *G*- and *Nv*-genes. Name of isolate, collection year(s), watershed(s), and homologous isolates are given. *=collected by us.

Isolate	Collection year of isolate (other year(s))	Watershed for isolate (others)	*P*-gene sequence (homologue(s))	*M*-gene sequence (homologue(s))
a. MI03GL	2003 (2006-9, 11)	L. St. Clair (all five Great Lakes)	a (c, e, f, g, h, i, j, l)	a (c, d, e, f, g, i)
b. vcG002	2008 (2005-9)	L. Michigan (Lakes Huron, Erie, Ontario)	b	b
c. vcG003	2007	L. Erie (L. Ontario)	(a)	(a)
d. vcG004	2006 (2008)	L. Erie	d	(a)
e. vcG005	2007	L. Huron	(a)	(a)
f. vcG006	2007	L. Huron	(a)	(a)
g. vcG007	2007	L. Erie	(a)	(a)
h. vcG008	2007	L. Michigan	(a)	h
i. vcG009	2008	L. Michigan	(a)	(a)
j. vcG010	2008	L. Michigan	(a)	j
k. LMB	2012	L. Erie	k	k
l. Drum	2012	L. Erie	(a)	l
Total *N* haplotypes			4	6

Table 3.2

Primer sequences for PCR and sequencing, references, annealing temperatures, and extension times.

Gene	Primer name	Reference	Sequence (5'-3')	Annealing temp. (^{o}C)	Extension time (min)
G	Gint1F	Thompson et al. (2011)	TCCCGTCAAGAGGCCAC	53	2:30
	Gint4R	"	TTCCAGGTGTTGTTTACCG	"	"
Nv	VHSvNv.se	Pore (2012)	ACGAATTCATGACGACCCAGTCGGCAC	57	1:30
	VHSvNv.as	"	ACGGTACCTGGGGGAGATTCGGAGCCA	"	"
	MI03GLNv.for	This study	GCACCCCTGTGAGACAGAAA	55	"
	MI03GLNv.rev	"	TGGGAGAAGGGGGAGGAG	"	"
P	VHSvPFor2	This study	CGCTGAGAGCTCACAATGAC	"	2:00
	VHSvRev2	"	GCCTTGATTGCCTTTGAGAC	"	"
M	VHSvM.se	Pore (2012)	ACGAATTCATGGCTCTATTCAAAAG	"	"
	VHSvM.as	"	ACGGTACCCCGGGGTCGGACAGAG	"	"

Table 3.3

Nucleotide substitutions, with base number and codon position, for unique haplotypes of the (a) *G*-, (b) *Nv*-, (c) *P*-, and (d) *M*-genes. The original MI03GL isolate is used as the baseline (top). .=nucleotide identical to MI03GL. *=described in this study.

(a)

		Nucleotide position (in reference to aligned GenBank sequence)													
		182	192	346	347	406	421	423	464	476	521	561	588	590	599
Codon Position		2	3	1	2	1	1	3	2	2	2	3	3	2	2
Haplotypes	GenBank Accession No.														
a. MI03GL	GQ385941	A	A	C	A	A	G	A	C	C	G	T	A	T	C
b. vcG002	EF564588	.	.	.	.	.	.	.	.	.	.	.	.	.	G
c. vcG003	HQ623435	.	.	.	.	.	.	.	.	.	A	.	.	.	.
d. vcG004	HQ623436	.	.	.	.	.	.	.	.	.	.	.	C	.	.
e. vcG005	HQ623437	G	.	.	.	.	.	.	.	.	.	.	.	.	.
f. vcG006	HQ623438	.	.	A	.	.	.	.	.	.	.	.	.	.	.

g. vcG007	HQ623439	.	.	.	.	.	A	.	.	.	.	.	.	A	.						
h. vcG008	HQ623440	.	G	.	.	.	.	.	.	.	.	.	.	.	.						
i. vcG009	HQ623441	.	.	.	.	G	.	G	.	.	.	.	.	.	.						
j. vcG010	HQ623442	.	.	.	G	.	.	.	.	.	.	.	.	.	G						
k. LMB*	XXXXXX	.	.	.	.	.	.	.	T	T	.	C	.	.	.						
l. Drum*	XXXXXX	.	.	.	.	.	.	.	T	.	.	.	.	.	.						

(b)

		Nucleotide position (in reference to aligned GenBank sequence)																			
		8	13	34	72	79	87	111	120	127	132	133	139	153	157	159	166	167	174	176	179
Codon Position		2	1	1	3	1	3	3	3	1	3	1	1	3	1	3	1	2	3	2	2
Haplotypes	GenBank Accession No.																				
a. MI03GL	GQ385941	T	C	T	A	T	A	C	A	T	T	A	T	T	T	T	T	T	T	T	G
b. vcG002*	XXXXXX	C	T	.	.	.	.	.	.	.	.	.	.	.	.	C	C	C	C	C	.
c. vcG003*	XXXXXX	.	T	.	.	.	.	.	.	C	C	.	C	.	C	.	.	.	.	.	.
d. vcG004*	XXXXXX	.	T	.	G	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.

e. vcG005*	XXXXXX	.	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
f. vcG006*	XXXXXX	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
g. vcG007*	XXXXXX	.	T	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.
h. vcG008*	XXXXXX	.	T	.	.	.	.	.	C	.	.	.	.	C	.	.	.	.	.	.	A
i. vcG009*	XXXXXX	.	T	.	.	G	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.
j. vcG010*	XXXXXX	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
k. LMB*	XXXXXX	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
l. Drum*	XXXXXX	T	.	C	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.

	194	199	210	213	218	224	233	243	247	251	254	270	276	279	288	303	310	313	323	324	332	337
Codon pos.	2	1	3	3	2	2	2	3	1	2	2	3	3	3	3	3	1	1	2	3	2	1
Haplotypes																						
a. MI03GL	T	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	A	T	T	T	A
b. vcG002*	.	.	.	.	.	.	.	.	.	C	C	.	C	.	.	.	.	.	C	.	.	.
c. vcG003*	C	C	.	C	C	C	C	C	C	.	C	C	.	C	C	C	C	.	C	C	C	.
d. vcG004*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
e. vcG005*	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

f. vcG006*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.
g. vcG007*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
h. vcG008*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
i. vcG009*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
j. vcG010*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
k. LMB*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
l. Drum*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	341	351	354	357	363	364
Codon pos.	2	3	3	3	3	1
Haplotypes						
a. MI03GL	T	C	T	G	T	T
b. vcG002*	C	.	C	A	C	C
c. vcG003*	C	.	C	A	C	C
d. vcG004*	.	.	C	A	C	C
e. vcG005*	.	.	C	A	C	C
f. vcG006*	.	.	C	A	C	C

g. vcG007*	.	.	C	A	C	C
h. vcG008*	.	.	C	A	C	C
i. vcG009*	.	.	C	A	C	C
j. vcG010*	.	.	C	A	C	C
k. LMB*	.	.	.	.	.	.
l. Drum*	.	A	.	.	.	.

(c)

		Nucleotide position (in reference to aligned GenBank sequence)					
		27	231	468	486	493	523
Codon Position		3	3	3	3	1	1
Haplotypes	GenBank Accession No.						
a. MI03GL	GQ385941	C	T	G	C	C	C
b. vcG002*	XXXXXX	.	.	T	T	A	T
d. vcG004*	XXXXXX	T	.	.	.	.	.
k. LMB*	XXXXXX	.	C	.	.	.	.

(d)

		Nucleotide position (in reference to aligned GenBank sequence)				
		111	153	329	372	543
Codon Position		3	3	2	3	3
Haplotypes	GenBank Accession No.					
a. MI03GL	GQ385941	G	T	A	T	A
b. vcG002*	XXXXXX	.	.	G	.	.
h. vcG008*	XXXXXX	T	.	.	.	.
j. vcG010*	XXXXXX	.	.	.	.	G
k. LMB*	XXXXXX	.	.	.	C	.
l. Drum*	XXXXXX	.	C	.	C	.

Table 3.4

Comparative genetic diversity values from 12 VHSv-IVb isolates for the *G*-, *Nv*-, *P*-, and *M*-genes.

	G	*Nv*	*P*	*M*
Sequence Diversity				
Sequence length (bp)	669	366	539	564
N of haplotypes (% of 12 isolates)	12 (100%)	12 (100%)	4 (30%)	6 (50%)
Haplotypic diversity (H_D)	1.00±0.01	1.00±0.01	0.56±0.04	0.68±0.04
N of homologous sequences (%)	0 (0%)	0 (0%)	8 (70%)	6 (50%)
Mutations				
N of substituted nucleotide sites (% of total bp)	14 (2.1%)	48 (13.1%)	6 (1.1%)	5 (0.9%)
Mutation rate: sub. site/yr (μ)	$2.1x10^{-4}$	$2.0x10^{-3}$	$1.5x10^{-4}$	$1.2x10^{-4}$*
N of 1st codon position substitutions (% of total)	3 (21%)	14 (29%)	2 (33%)	0 (0%)
N of 2nd codon position substitutions (% of total)	7 (50%)	13 (27%)	0 (0%)	1 (20%)
N of 3rd codon position substitutions (% of total)	4 (29%)	21 (44%)	4 (67%)	4 (80%)
N synonymous changes (% of substituted sites)	6 (43%)	19 (40%)	3 (50%)	3 (60%)
N non-synonymous/amino acid changes (% of substituted sites)	8 (57%)	29 (60%)	3 (50%)	2 (40%)
Divergence				
Pairwise *N* of nucleotide substitutions between haplotypes: mean±SE (range)	2.8±0.4 (1-7)	10.7±2.3 (1-31)	2.5±0.7 (1-4)	1.9±0.3 (1-3)
Pairwise (*p*) distances between haplotypes: mean±SE (range)	0.004±0.002 (0.001-0.010)	0.028±0.008 (0.003-0.085)	0.006±0.003 (0.002-0.009)	0.003±0.002 (0.002-0.005)

Fig. 3-1 Diagram of the VHSv genome, showing its six proteins and bullet-shaped virion structure (modified, with permission, from Pore, 2012).

Fig. 3-2 Maps showing the distribution of VHSv-IVb isolates and unique haplotypes (lettered according to Table 2) across the North American Laurentian Great Lakes for the (a) *G*- (b) *Nv*-, (c) *P*-, and (d) *M*-genes. Circles=isolates sequenced by us in the present study. Squares=individuals sequenced by other researchers and not available to our study.

Fig. 3-3 Numbers of nucleotide substitutions (△, ...), transitions (●, ——), and transversions (○, – –), versus uncorrected pairwise (*p*-) distances of the combined sequence haplotypes for: (a) the combined sequences for the four genes and for the (b) *G*-, (c) *Nv*-, (d) *P*-, and (e) *M*-genes. Results from ANCOVA between the relationships of transitional and transversional substitutions are given under the gene heading. Regression lines are omitted when R^2<0.50.

Fig. 3-4 Numbers of substitutions (△, ...), synonymous (■, ——), and non-synonymous changes (□, – –), versus uncorrected pairwise (*p*-) distance of the combined sequence haplotypes for: (a) the combined sequences for the four genes and for the (b) *G*-, (c) *Nv*-, (d) *P*-, and (e) *M*-genes. ANCOVA results between synonymous and non-synonymous changes are located under each heading. Regression lines are omitted when R^2<0.50.

Fig. 3-5 Haplotype networks of VHSv-IVb variants from *TCS* 1.21 (Clement et al.,

2000) for: (a) combined sequences of the four genes and for the (b) *G*-, (c) *Nv*-, (d) *P*-, and (e) *M*-genes. Circles are sized according to population frequency of the haplotype and lettered according to Table 3.2. Samples in squares were unavailable for sequencing by us (thus there are no data for the other genes). Parentheses enclose the total number of sequences for haplotypes having >1 individual (see Appendix 3.1). Lines denote a single mutational step between haplotypes; small, unlabeled black circles represent hypothesized haplotypes. Collection location(s) and year(s) are listed beside each haplotype.

Fig. 3-6 Phylogenetic trees of VHSv haplotypes based on (a) combined sequences for the four genes and for the (b) *G*-, (c) *Nv*-, (d) *P*-, and (e) *M*- genes, from maximum likelihood and Bayesian analyses. Values above nodes=2000 bootstrap pseudoreplicates/Bayesian posterior probabilities. Estimated divergence times (years) are in parentheses. Trees are rooted to VHSv strain I, II, and III sequences.

Fig. 3-1

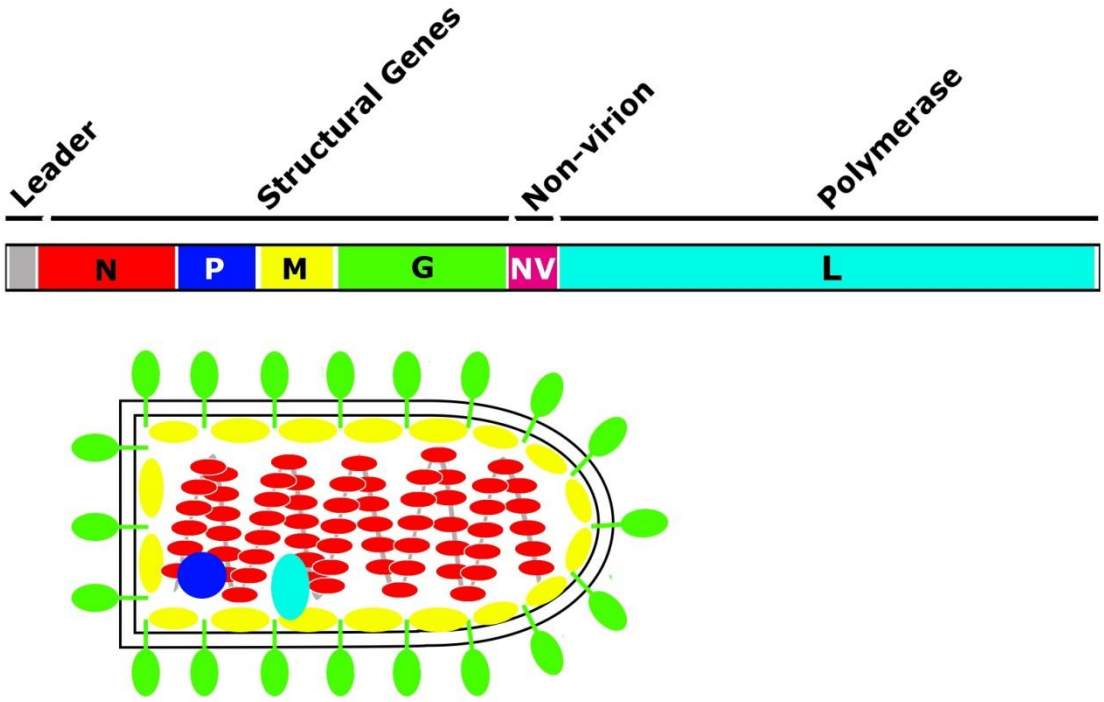


Fig. 3-2

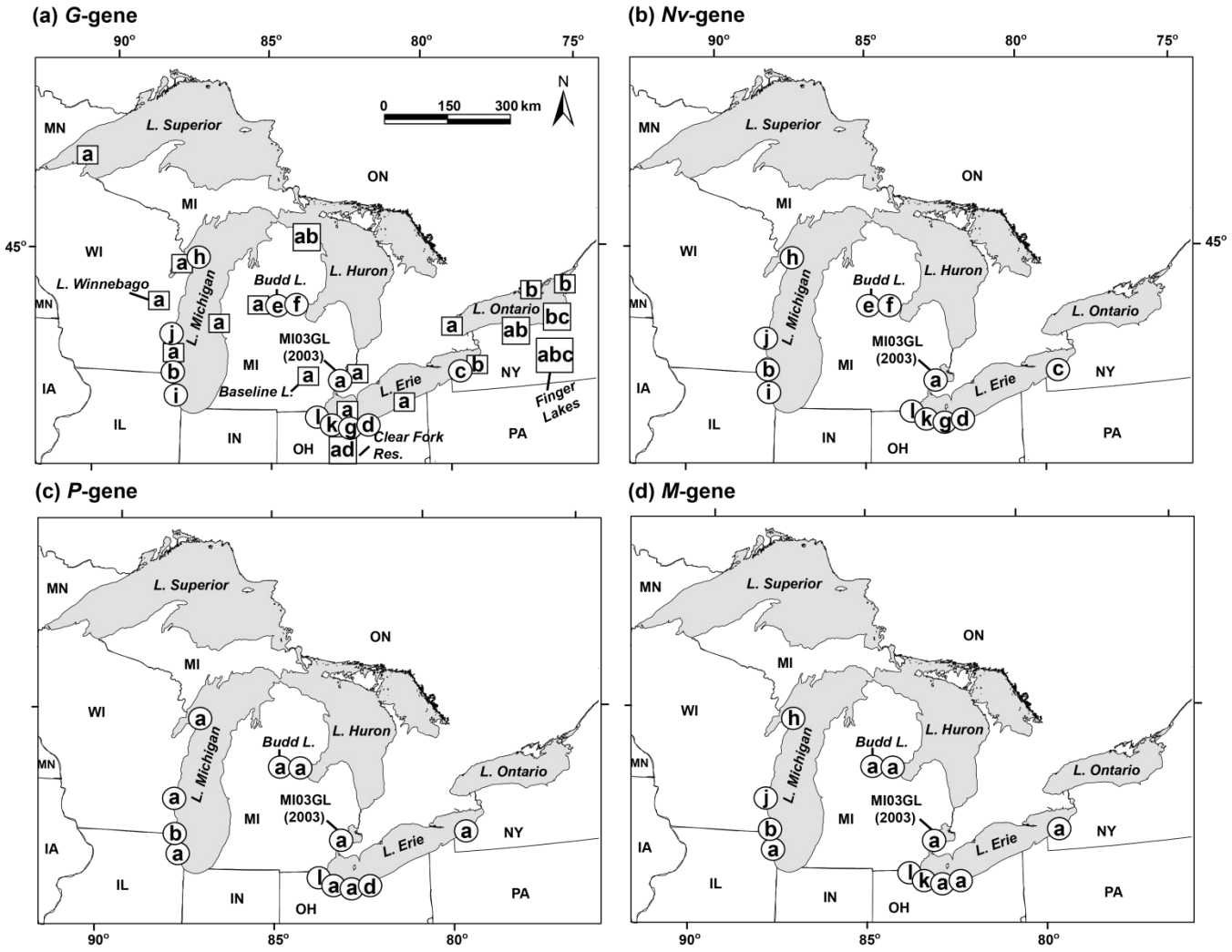

(a) G-gene
(b) Nv-gene
(c) P-gene
(d) M-gene
90°
85°
80°
75°
45°
0 150 300 km
N
MN
WI
MI
ON
IA
IL
IN
OH
NY
PA
L. Superior
L. Michigan
L. Huron
Budd L.
L. Ontario
L. Erie
L. Winnebago
Baseline L.
MI03GL (2003)
Clear Fork Res.
Finger Lakes

Fig. 3-3

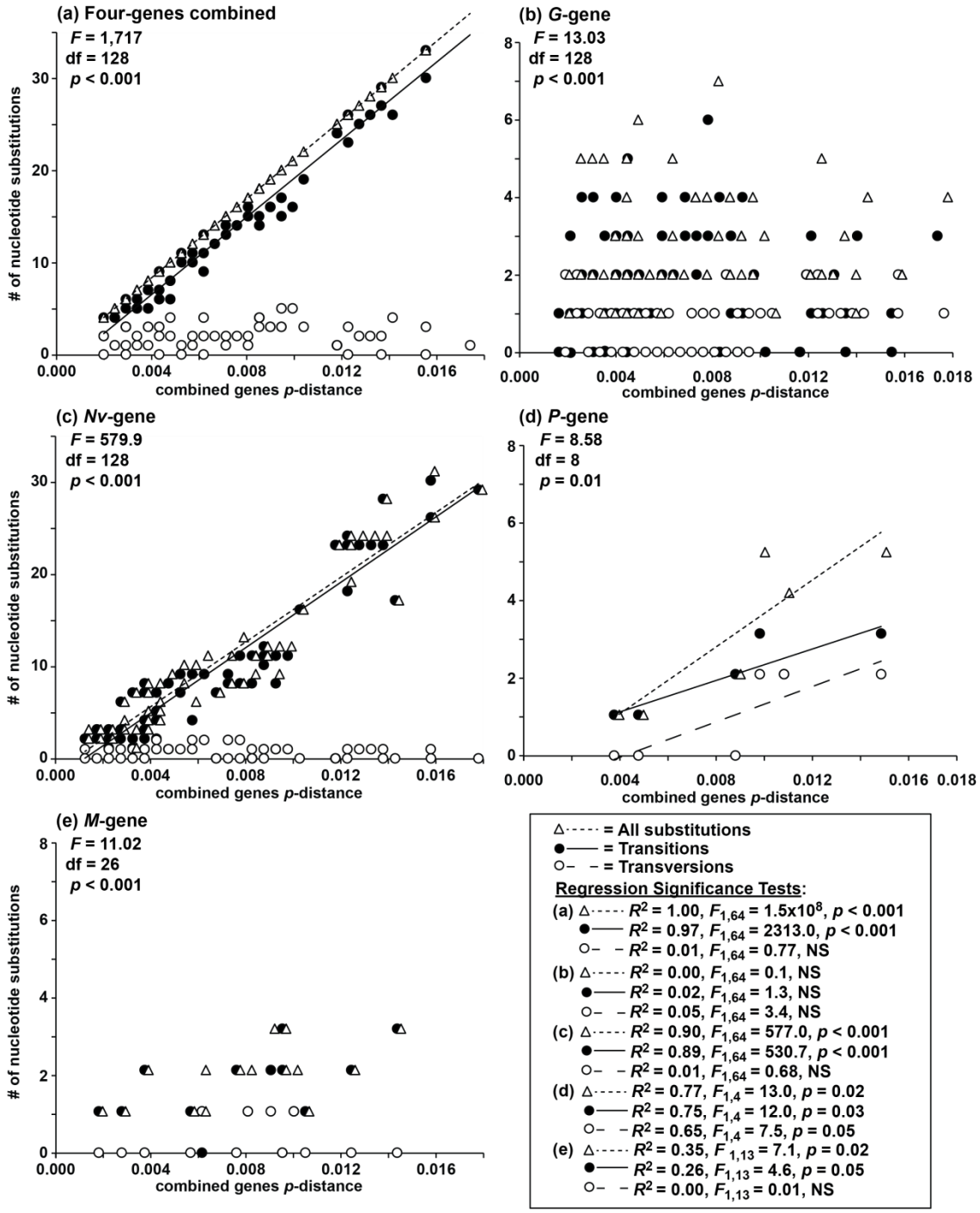

(a) Four-genes combined
F = 1,717
df = 128
p < 0.001
(b) G-gene
F = 13.03
df = 128
p < 0.001
(c) Nv-gene
F = 579.9
df = 128
p < 0.001
(d) P-gene
F = 8.58
df = 8
p = 0.01
(e) M-gene
F = 11.02
df = 26
p < 0.001
# of nucleotide substitutions
combined genes p-distance
△····· = All substitutions
●—— = Transitions
○– – = Transversions
Regression Significance Tests:
(a) △····· R² = 1.00, F1,64 = 1.5x10⁸, p < 0.001
●—— R² = 0.97, F1,64 = 2313.0, p < 0.001
○– – R² = 0.01, F1,64 = 0.77, NS
(b) △····· R² = 0.00, F1,64 = 0.1, NS
●—— R² = 0.02, F1,64 = 1.3, NS
○– – R² = 0.05, F1,64 = 3.4, NS
(c) △····· R² = 0.90, F1,64 = 577.0, p < 0.001
●—— R² = 0.89, F1,64 = 530.7, p < 0.001
○– – R² = 0.01, F1,64 = 0.68, NS
(d) △····· R² = 0.77, F1,4 = 13.0, p = 0.02
●—— R² = 0.75, F1,4 = 12.0, p = 0.03
○– – R² = 0.65, F1,4 = 7.5, p = 0.05
(e) △····· R² = 0.35, F 1,13 = 7.1, p = 0.02
●—— R² = 0.26, F1,13 = 4.6, p = 0.05
○– – R² = 0.00, F1,13 = 0.01, NS

Fig. 3-4

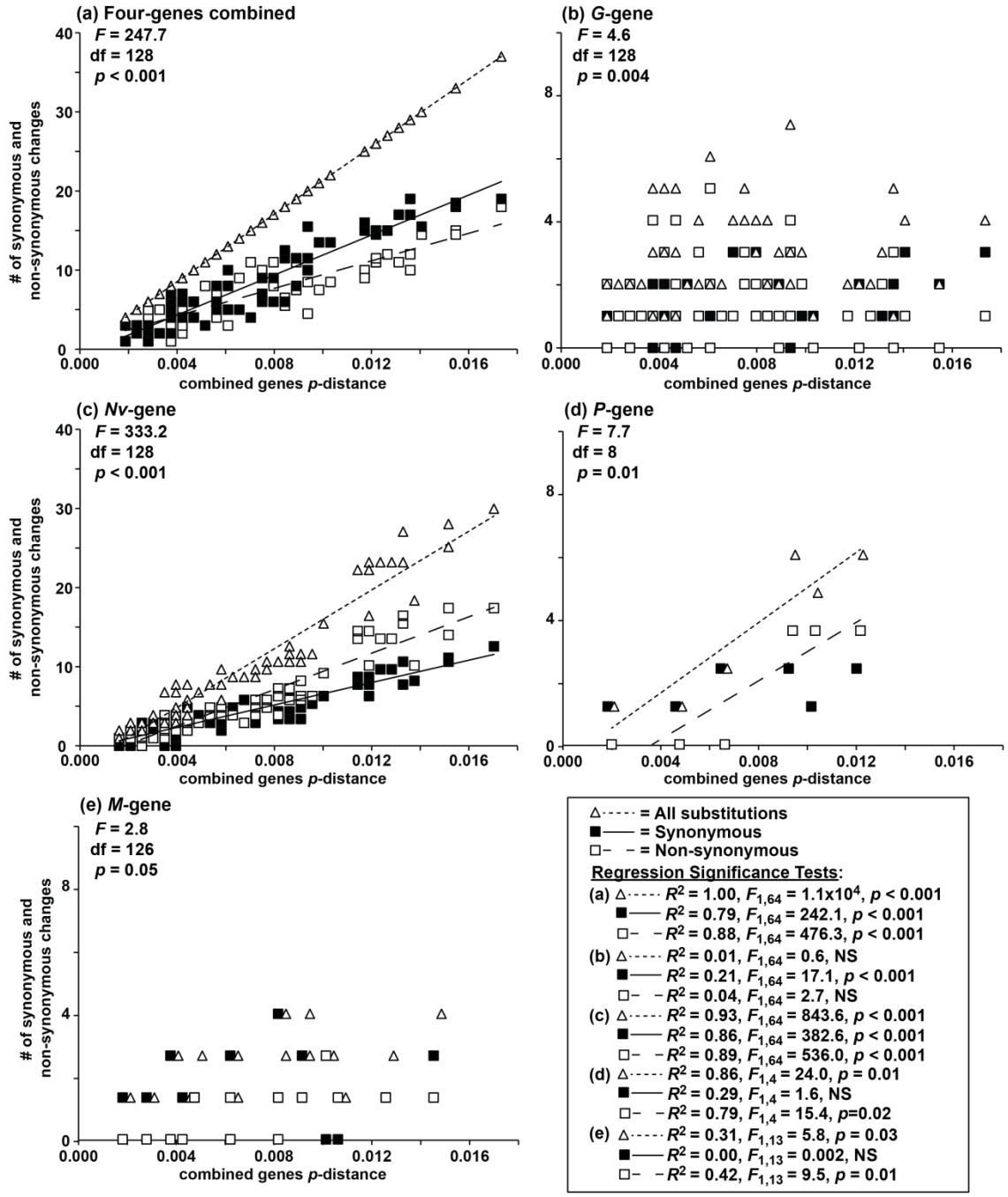

(a) Four-genes combined
F = 247.7
df = 128
p < 0.001
(b) G-gene
F = 4.6
df = 128
p = 0.004
(c) Nv-gene
F = 333.2
df = 128
p < 0.001
(d) P-gene
F = 7.7
df = 8
p = 0.01
(e) M-gene
F = 2.8
df = 126
p = 0.05
# of synonymous and non-synonymous changes
combined genes p-distance
0.000
0.004
0.008
0.012
0.016
△····· = All substitutions
■—— = Synonymous
□– – = Non-synonymous
Regression Significance Tests:
(a) △····· R² = 1.00, F1,64 = 1.1x10⁴, p < 0.001
■—— R² = 0.79, F1,64 = 242.1, p < 0.001
□– – R² = 0.88, F1,64 = 476.3, p < 0.001
(b) △····· R² = 0.01, F1,64 = 0.6, NS
■—— R² = 0.21, F1,64 = 17.1, p < 0.001
□– – R² = 0.04, F1,64 = 2.7, NS
(c) △····· R² = 0.93, F1,64 = 843.6, p < 0.001
■—— R² = 0.86, F1,64 = 382.6, p < 0.001
□– – R² = 0.89, F1,64 = 536.0, p < 0.001
(d) △····· R² = 0.86, F1,4 = 24.0, p = 0.01
■—— R² = 0.29, F1,4 = 1.6, NS
□– – R² = 0.79, F1,4 = 15.4, p=0.02
(e) △····· R² = 0.31, F1,13 = 5.8, p = 0.03
■—— R² = 0.00, F1,13 = 0.002, NS
□– – R² = 0.42, F1,13 = 9.5, p = 0.01

Fig. 3-5

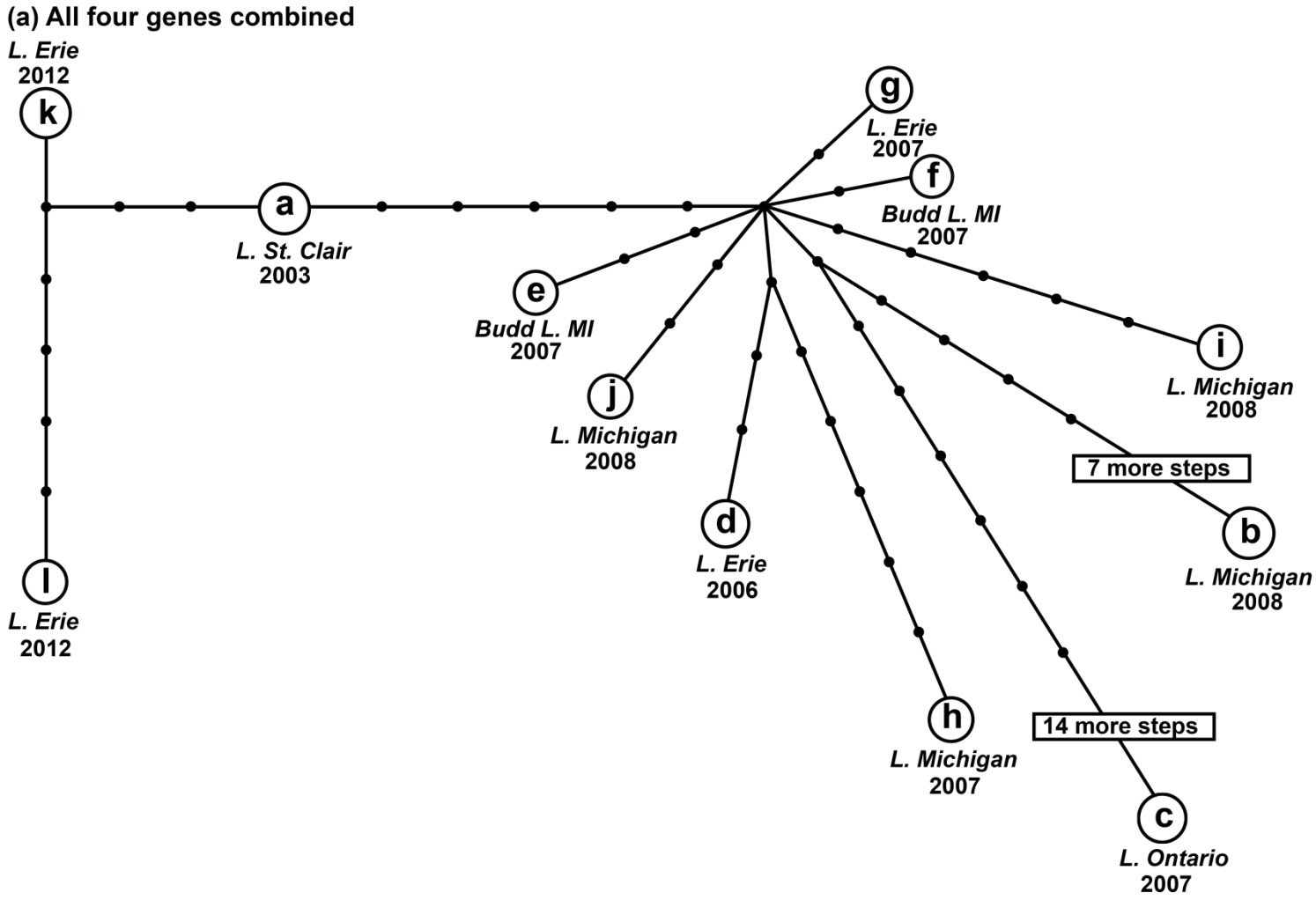

(a) All four genes combined
L. Erie
2012
k
a
L. St. Clair
2003
l
L. Erie
2012
g
L. Erie
2007
f
Budd L. MI
2007
e
Budd L. MI
2007
j
L. Michigan
2008
d
L. Erie
2006
i
L. Michigan
2008
7 more steps
b
L. Michigan
2008
h
L. Michigan
2007
14 more steps
c
L. Ontario
2007

Fig. 3-5

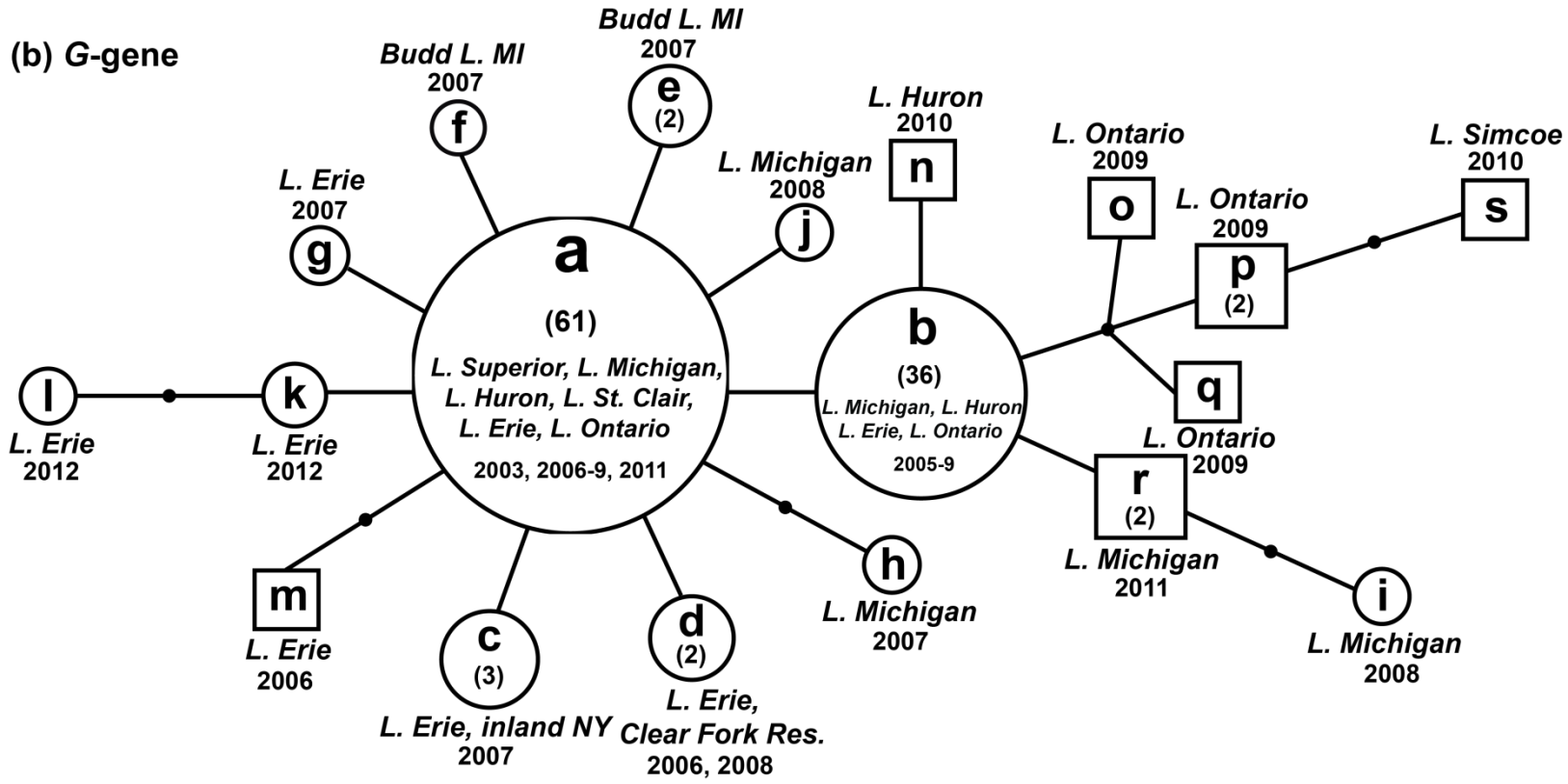

(b) G-gene
a
(61)
L. Superior, L. Michigan,
L. Huron, L. St. Clair,
L. Erie, L. Ontario
2003, 2006-9, 2011
b
(36)
L. Michigan, L. Huron
L. Erie, L. Ontario
2005-9
Budd L. MI
2007
f
Budd L. MI
2007
e
(2)
L. Erie
2007
g
L. Michigan
2008
j
L. Huron
2010
n
L. Ontario
2009
o
L. Ontario
2009
p
(2)
L. Simcoe
2010
s
q
L. Ontario
2009
r
(2)
L. Michigan
2011
i
L. Michigan
2008
l
L. Erie
2012
k
L. Erie
2012
m
L. Erie
2006
c
(3)
L. Erie, inland NY
2007
d
(2)
L. Erie,
Clear Fork Res.
2006, 2008
h
L. Michigan
2007

Fig. 3-5

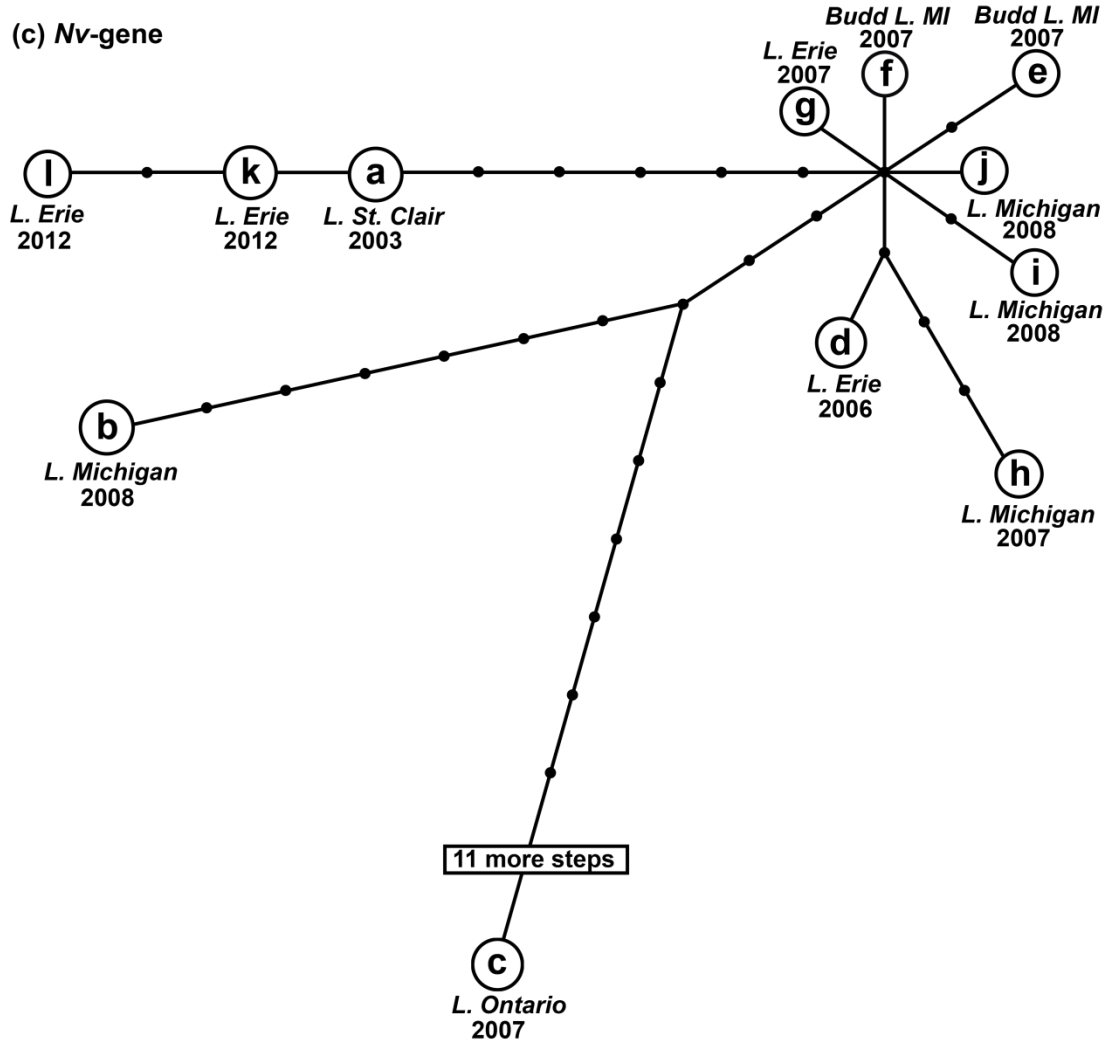

(c) Nv-gene
l
L. Erie
2012
k
L. Erie
2012
a
L. St. Clair
2003
g
L. Erie
2007
f
Budd L. MI
2007
e
Budd L. MI
2007
j
L. Michigan
2008
i
L. Michigan
2008
d
L. Erie
2006
h
L. Michigan
2007
b
L. Michigan
2008
11 more steps
c
L. Ontario
2007

Fig. 3-5

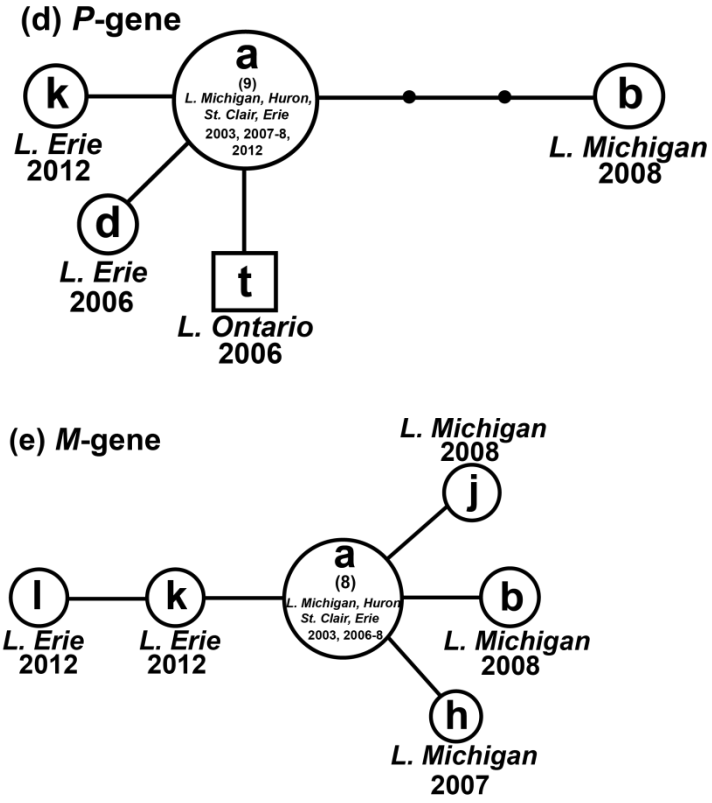

(d) P-gene
a
(9)
L. Michigan, Huron, St. Clair, Erie
2003, 2007-8, 2012
k
L. Erie
2012
b
L. Michigan
2008
d
L. Erie
2006
t
L. Ontario
2006
(e) M-gene
L. Michigan
2008
j
a
(8)
L. Michigan, Huron, St. Clair, Erie
2003, 2006-8
l
L. Erie
2012
k
L. Erie
2012
b
L. Michigan
2008
h
L. Michigan
2007

Fig. 3-6

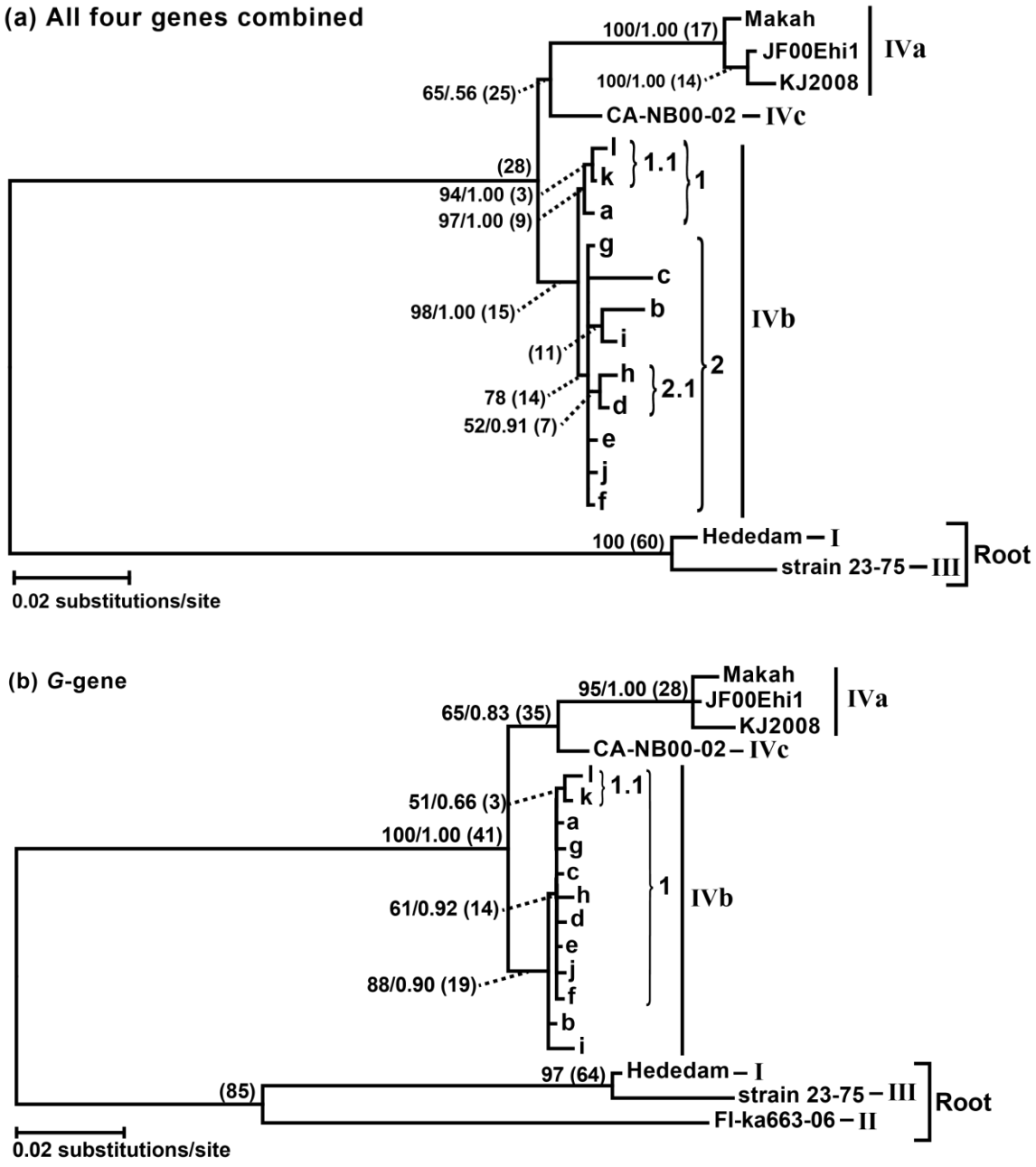

(a) All four genes combined
100/1.00 (17)
Makah
JF00Ehi1
KJ2008
IVa
100/1.00 (14)
65/.56 (25)
CA-NB00-02 – IVc
(28)
l
k
1.1
1
94/1.00 (3)
97/1.00 (9)
a
g
c
98/1.00 (15)
b
i
IVb
(11)
h
2.1
2
78 (14)
d
52/0.91 (7)
e
j
f
100 (60)
Hededam – I
strain 23-75 – III
Root
0.02 substitutions/site
(b) G-gene
95/1.00 (28)
Makah
JF00Ehi1
KJ2008
IVa
65/0.83 (35)
CA-NB00-02 – IVc
l
k
1.1
51/0.66 (3)
a
100/1.00 (41)
g
c
h
1
IVb
61/0.92 (14)
d
e
j
88/0.90 (19)
f
b
i
97 (64)
Hededam – I
(85)
strain 23-75 – III
Root
Fl-ka663-06 – II
0.02 substitutions/site

Fig. 3-6

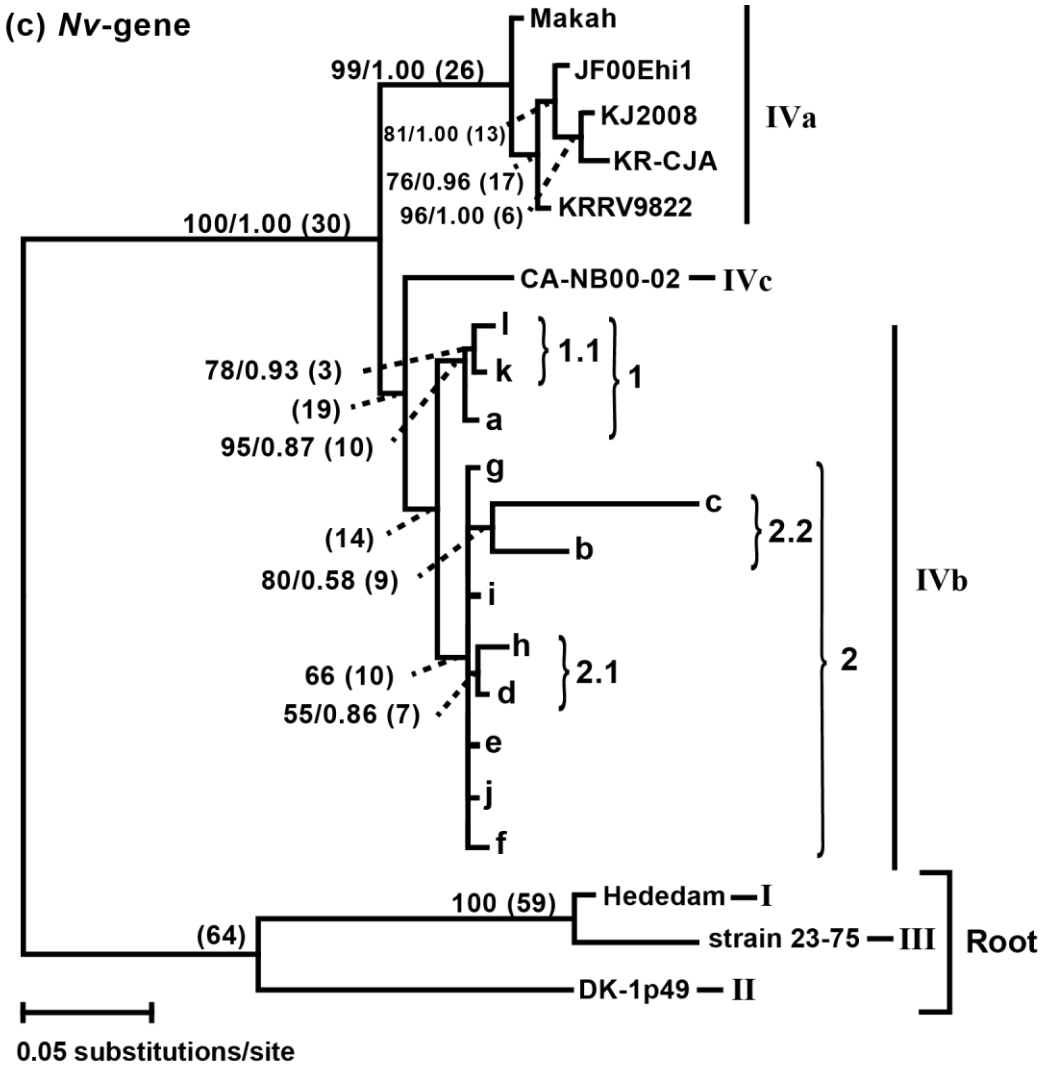
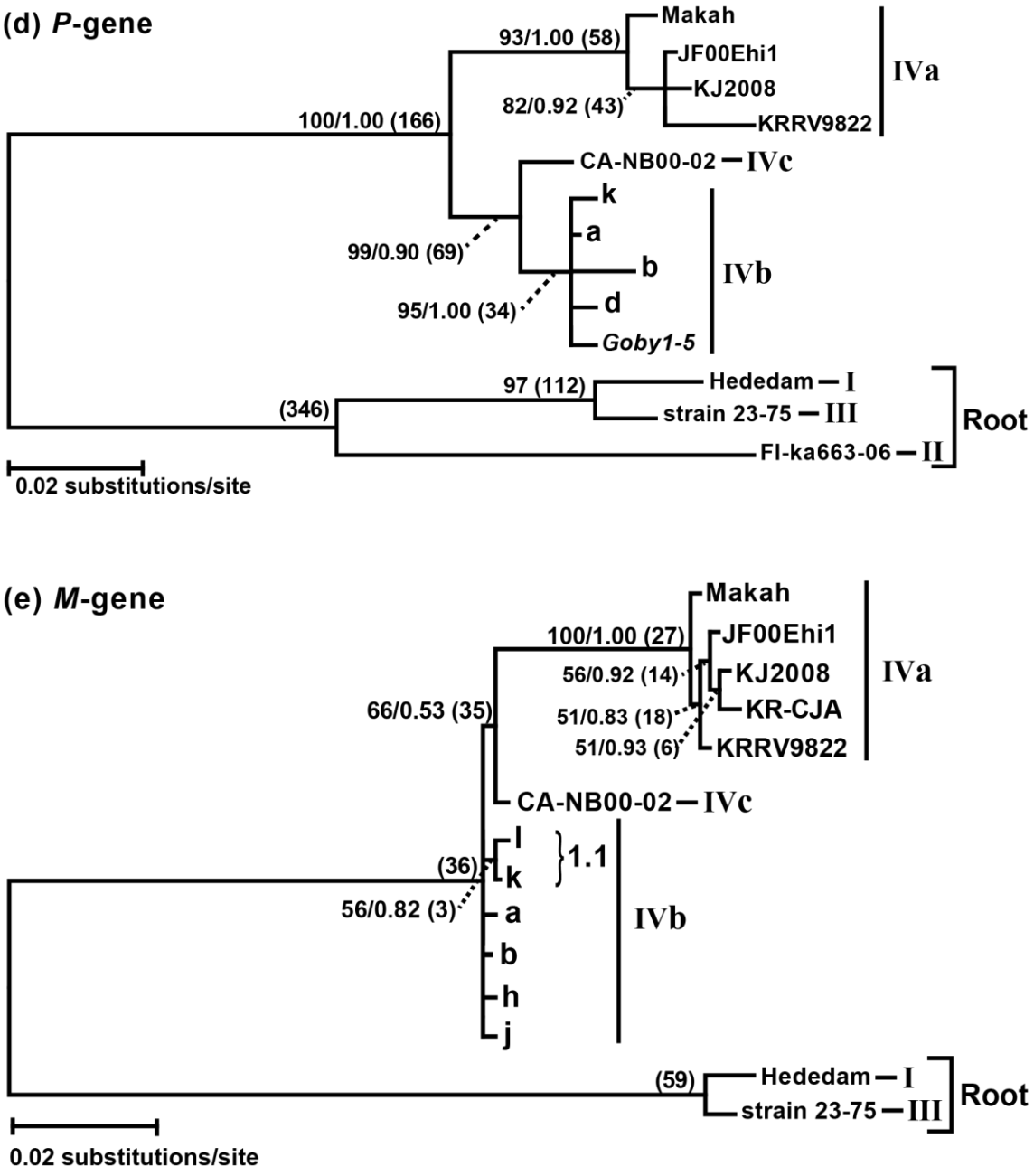

(c) Nv-gene
Makah
99/1.00 (26)
JF00Ehi1
KJ2008
81/1.00 (13)
KR-CJA
76/0.96 (17)
KRRV9822
96/1.00 (6)
IVa
100/1.00 (30)
CA-NB00-02
IVc
l
1.1
78/0.93 (3)
k
1
(19)
a
95/0.87 (10)
g
c
2.2
(14)
b
80/0.58 (9)
i
IVb
h
2.1
66 (10)
d
2
55/0.86 (7)
e
j
f
100 (59)
Hededam
I
(64)
strain 23-75
III
Root
DK-1p49
II
0.05 substitutions/site

Fig. 3-6


(d) P-gene
Makah
93/1.00 (58)
JF00Ehi1
KJ2008
82/0.92 (43)
KRRV9822
IVa
100/1.00 (166)
CA-NB00-02
IVc
k
a
99/0.90 (69)
b
IVb
95/1.00 (34)
d
Goby1-5
97 (112)
Hededam
I
(346)
strain 23-75
III
Root
FI-ka663-06
II
0.02 substitutions/site
(e) M-gene
Makah
100/1.00 (27)
JF00Ehi1
56/0.92 (14)
KJ2008
IVa
66/0.53 (35)
51/0.83 (18)
KR-CJA
51/0.93 (6)
KRRV9822
CA-NB00-02
IVc
l
1.1
(36)
k
56/0.82 (3)
a
IVb
b
h
j
(59)
Hededam
I
Root
strain 23-75
III
0.02 substitutions/site

Appendix 3.1

VHSv *G*-, *Nv*-, *P*-, and *M*-gene sequences per strain, host species, isolate, GenBank accession number, homologous sequences, reference, and locality. *=sequenced by us in this study.

Gene	Strain	Host Species	Isolate	Accession No.	Homologous Sequences: Isolate (Accession No.)	Reference(s)	Locality	Lat.	Long.
G	**I**	*Oncorhynchus mykiss*	Hededam	Z93412	__	Stone et al., (1997)	Spjarup Hededamme, DNK	56.648	9.271
	II	*Clupea harengus*	ka663-06	HQ112247	ka664_06 (HQ112231), ka662_06 (HQ112230)	Gadd et al., (2011)	Archipelago S., FIN	60.290	21.290
	III	*Salmo trutta*	strain 23-75	FN665788	FR-2375 (AY546617)	Biacchesi et al., (2010), Einer-Jensen et al., (2004)	FRA	__	__
	IVa	*Clupea harengus*	BC93-390 BC96-265-3 BC96-265-6	KC117214, KC117217-8	__	Garver et al., (2013)	Campbell R., BC, CAN	50.024	-125.248
		" "	ME03	DQ401192	__	Elsayed et al., (2006)	N. Pacific, USA	47.622	-122.638
		C. harengus pallasii	BC98-249,	KC117219,	__	Garver et al., (2013),	Salt Spring Is., BC, CAN;	48.815,	-123.508,
			Campbell River248 BC 1993	U88051		Benmansour et al., (1997)	Campbell R., BC, CAN	50.024	-125.248
		" "	BC00-LF	KC117220	__	Garver et al., (2013)	Nanoose Harb., BC, CAN	49.161	-123.929
		" "	BC00-397	KC117221	__	" "	Hectate Strt., BC, CAN	53.193	-130.821
		" "	BC02-47-21	KC117226	__	" "	Stopper Is., Barkley Sound, BC, CAN	48.983	-125.353
		" "	BC02-235-2	KC117227	__	" "	Hardy Bay, BC, CAN	50.712	-127.491
		" "	BC05-014-2, BC05-014-7	KC117233-4	__	" "	Clayoquot Sound, BC, CAN	49.250	-126.000
		" "	BC07-13-2	KC117245	__	" "	Arrow Pass, BC, CAN	48.450	-123.330

“ “	BC09-31-8, BC09-31-2	KC117247-8	__	“ “	Port Hardy, BC, CAN	50.717	-127.500
“ “	WA-93	DQ473303	__	Hedrick et al., (2003)	Elliott Bay, WA, USA	47.622	-122.638
“ “	U88052,	U88052,	__	Benmansour et al., (1997), Stone et al., (1997)	Prince William Sound, AK, USA	61.622	-146.863
	AK93#1,	Z93430					
“ “	BC93372	DQ401186	__	Elsayed et al., (2006)	Prince Rupert Harb., BC, CAN	54.301	-130.332
“ “	BC99010,	DQ401194,	__	“ “	Gilford Is., BC, CAN;	50.717,	-126.369,
	BC99001	DQ401195			Beaver Cove, Telegraph Ck., BC, CAN	50.532	-126.856
Cymatogaster aggregata	BC-sp-02	DQ473301	__	Hedrick et al., (2003)	NW Vancouver Is., BC, CAN	49.233	-123.100
“ “	BC02-41-14	KC117225	__	Garver et al., (2013)	Fair Harb., Kyuquot Sound, BC, CAN	50.066	-127.129
Gadus macrocephalus	AK’93, NA-6, NA-7	Z93429, Z93424-5,	__	Stone et al., (1997),	Prince William Sound, AK, USA	61.622	-146.863
	AK-90	DQ473302		Hedrick et al., (2003)			
“ “	Clearwater WA 1991	U88050	__	Benmansour et al. (1997)	WA, USA	__	__
Merluccius productus	BC11-191	KC117251	__	Garver et al., (2013)	Bedwell Sound, BC, CAN	50.124	-127.129
“ “	WA91 Clearwater	DQ401189	__	Elsayed et al., (2006)	WA, USA	__	__
Oncorhynchus kisuth	Makah,	U28747, Z93421,	__	Benmansour et al., (1997),	Makah National Fish Hatchery, Neah Bay, WA, USA;	48.290,	-124.650,
	Elok,	Z93422		Stone et al., (1997)	Elokomin River, WA, USA;	46.224,	-123.330,
	NA-5,	Z93423,			Bogachiel R., WA, USA;	47.902,	-124.195,
	NA-8	Z93426			Clearwater R., WA, USA	47.733	-124.015
O. mykiss	RtGw11, RtGw10, RtGw5,	HQ687070-71, HQ687073	RtGw8(HQ687072)	Suebsing and Kim (unpub.)	S. KOR	__	__
O. tshawytscha	BC301-1A, BC283-1B, BC313-1A	KC117235-7	__	Garver et al., (2013)	W. Coast Vancouver Is., BC, CAN	49.900	-125.170

Paralichthys olivaceus	AY167587	AY167587	__	Kim and Park (unpub.)	S. KOR	__	__
" "	FWando08	GU265811	__	Oh et al., (unpub.)	S. KOR	__	__
" "	KR-CJA	JQ651388	KR-YGH (JQ651393)	Lee et al., (unpub.)	Wando, KOR, Jeju, KOR	34.389, 33.430	126.702, 126.546
" "	KRRV9822	AB179621	__	Byon et al., (unpub.)	S. of JPN, JPN	32.000	132.000
" "	OfGn	HQ687076	PhGn(HQ687075), PcGn(HQ687074)	Suebsing and Kim (unpub.)	S. KOR	__	__
" "	JP99Obama25, #25,	DQ401191, AB060725,	__	Elsayed et al., (2006), Nishizawa et al., (unpub.), Ito et al., (unpub.)	Obama coastal area of the Wakasa Bay, JPN;	35.500,	135.700,
	JF00Ehi1	AB490792			Ehime, JPN	33.750	132.600
" "	KJ2008	JF792424	FYeosu05 (FJ811901), FWando05 (FJ811900), FJeju05(FJ811902), FYG08(GU265812)	Kim and Kim (unpub.), Kim et al., (unpub.), Oh et al., (unpub.)	S. KOR	__	__
" "	GCVP-01, GCVP-03, GCVP-07, GCVP-09, GCVP-12, GCVP-18, GCVP-20, GCVP-27	JQ952775-82	__	Han et al., (unpub.)	" "	" "	" "
Salmo salar	BC95-297	KC117215	__	Garver et al., (2013)	Campbell R., BC, CAN	50.024	-125.248
" "	BC02-03, BC07-286-10, BC07-15-7	KC117223, KC117242, KC117244	__	" "	Arrow Pass, BC, CAN	48.450,	-123.330,
" "	BC04-028-1, BC04-040, BC05-011	KC117230-2	__	" "	Clayoquot Sound, BC, CAN	49.250	-126.000
" "	BC06-37-2	KC117241	__	" "	Bedwell Sound, BC, CAN	50.124	-127.129
" "	BC08-2816-1	KC117246	__	" "	Port Hardy, BC, CAN	50.717	-127.50
" "	BC10-42-13	KC117249	__	" "	Barkley Sound, BC, CAN	48.849	-125.386
Sardinops sagax	BC02-28-6, BC02-41-9	KC117222, KC117224	__	" "	Fair Harb., Kyuquot Sound, BC, CAN	50.066	-127.129

	" "	BC02-232-1	KC117228	__	" "	Quatsino Sound, BC, CAN	50.417	-128.000
	" "	BC02-229	KC117229	__	" "	Hardy Bay, BC, CAN	__	__
	" "	BC06-089-1, BC-06-089-4	KC117239-40	__	" "	Rennell Sound, BC, CAN	53.417	-132.750
	" "	BC07-21-2	KC117243	__	" "	Clayoquot Sound, BC, CAN	49.250	-126.000
	" "	BC05-197	KC117238	__	" "	Kitimat, BC, CAN	54.063	-128.578
	" "	BC98250,	DQ401187,	__	Elsayed et al., (2006)	Salt Spring Is., BC, CAN;	48.815,	-123.508,
		BC99292	DQ401188			S. Vancouver Is., BC, CAN	49.827	-125.620
	" "	CMLs02	DQ473300	__	Hedrick et al., (2003)	Moss Landing, CA, USA	36.804	-121.786
	" "	CMAs02	DQ473299	__	" "	Malibu, CA, USA	34.030	-118.780
	" "	CS01	DQ473298	__	" "	Los Angeles, CA, USA	34.052	-118.242
	" "	BC-s-99	DQ473296	__	" "	Queen Charlotte, BC, CAN	53.255	-132.087
	" "	Oe01	DQ473297	__	" "	Sandy R., OR, USA	45.401	-122.228
IVb	*Ambloplites rupestris*	TAVgr08-03 (vcG009)	HQ623441	__	Thompson et al., (2011)	N. Point Marina, L. MI, IL, USA	42.499	-90.689
	Aplodinotus grunniens	U13653-1 (vcG002)	HQ453209		Groocock et al., (2007),	Bay of Quinte, L. Ont., CAN;	43.968,	-77.629,
				RG06(EF564588), U13653-2,	Thompson et al., (2011)	St. Lawrence R., NY, USA;		
				TAVgr06-02, TAVgr06-27, TAVgr08-04,			44.250,	-75.000,
				TAVgr06-19,			44.327,	-75.937,
				TAVgr06-03, TAVgr06-04,			44.127,	-76.333,
				TAVgr06-31,			44.257,	-76.134,
				TAVgr06-32, TAVgr06-33,			44.323,	-76.014,

			TAVgr06-34,			44.248,	-76.014,
			TAVgr06-35,			44.254,	-76.014,
			TAVgr06-36,			44.187,	-76.014,
			TAVgr06-39,				
			TAVgr06-37,			44.175,	-76.964,
			TAVgr06-38,			44.323,	-76.014,
			TAVgr06-24,			44.172,	-76.247,
			TAVgr06-20,			44.323,	-75.935,
			TAVgr06-25,				
			TAVgr06-21,			44.242,	-76.098,
			TAVgr06-26,			44.187,	-76.225,
			TAVgr06-23,				
			TAVgr06-40,			44.268,	-76.014,
			TAVgr06-22,			44.254,	-76.150,
			TAVgr08-07,			43.340,	-75,910,
			TAVgr07-08,		Dunkirk Harb., L. Erie, NY, USA; L.	42.490,	-79.338,
			TAVgr08-08,		Huron, USA; L.	45.625,	-84.468,
			TAVgr08-02,		Michigan, WI, USA;	42.799,	-87.760,
			TAVgr08-06,		Fairhaven State Park, L. Ont., NY, USA;	43.350,	-76.690,
			TAVgr08-05,		Oswego, L. Ont., NY, USA;	43.450,	-76.510
			TAVgr06-30,		Tibbetts Ck., L. Ont. USA;	44.116	-76.333,
			TAVgr06-01,		W. of Rochester, L. Ont., USA;	43.216	-77.633,
			TAVgr07-18, TAVgr07-19		L. Skaneateles, NY, USA	42.490	-79.338
"	"	*DRUM	XXXXXX	__	This study	Sandusky B., L. Erie, OH, USA	41.453	-82.726
Cyprinus carpio		TAVgr07-12 (vcG003)	HQ623435		Thompson et al., (2011)	Dunkirk Harb., L. Erie, NY, USA;	42.490,	-79.338,
				TAVgr07-17,		Little Salmon R., NY, USA;	43.459,	-76.228,
				TAVgr07-13		Skaneateles L., NY, USA	42.950	-76.230
"	"	r. TAVgr11-02 (vcG028)	__	vcG028	Kurath et al., (unpub.)	Menomonee R. Canal, WI, USA	43.030	-87.915

Esox masquinongy	MI03GL	GQ385941	DQ401193,	Elsayed et al., (2006), Thompson et al., (2011)	L. St. Clair, MI, USA;	42.390,	-82.911,
			TAVgr05-01, 0601FD, 0602SB, 0603BG,				
			TAVgr06-48,			42.343,	-82,902,
			TAVgr06-51,				
			TAVgr06-52,				
			TAVgr06-50,				
			TAVgr06-47,			42,634,	-82.777,
			TAVgr06-49,				
			TAVgr09-03,			42.631,	-82,765,
			TAVgr09-04,				
			TAVgr09-10,			42.615,	-82,757,
			TAVgr09-11,			42.475,	-82.879,
			TAVgr07-09,		Dunkirk Harb.,	42.490,	-79.338,
			TAVgr09-05,		L. Erie, NY, USA; L. Erie, USA;	—	—
			TAVgr06-07,		L. Erie, OH, USA	41.755,	-81.286,
			TAVgr06-08,				
			TAVgr06-09,				
			TAVgr06-10,				
			TAVgr06-11,				
			TAVgr06-15,				
			TAVgr06-53,				
			TAVgr07-21,				
			TAVgr07-22,				
			TAVgr08-09,				
			TAVgr08-10,				
			TAVgr08-11,				
			TAVgr06-05,		Sandusky Bay, L. Erie, OH, USA;	41.474,	-82.703,
			TAVgr06-06,				
			TAVgr06-12,		W. L. Erie, OH, USA;	41.492,	-82.667,
			TAVgr06-13,				
			TAVgr06-14,				
			TAVgr06-17,				
			TAVgr06-18,				
			TAVgr06-46,		Cheboygan Bay, L. Huron, MI, USA;	45.718,	-84.374,
			TAVgr06-44,		Swan R., L. Huron, MI, USA;	45.502,	-83.783,

			TAVgr06-43, TAVgr06-45, TAVgr09-01, TAVgr09-02, TAVgr09-12, OMNR 5577, OMNR 5583, OMNR#5579,		Thunder Bay, L. Huron, MI, USA;	45.050,	-83.200,
			TAVgr06-28,		L. Mich. MI, USA;	43.600,	-86,916,
			TAVgr07-06, TAVgr07-07,		Sturgeon Bay, L. Mich., WI, USA;	44.860,	-87,393,
			TAVgr06-29,		Hamilton Harb. L. Ont., ON, CAN;	44.514,	-87.831,
			TAVgr10-01,		Thames R., L. ON, ON, CAN;	43.295,	-79.772,
			TAVgr09-09,		Irondequoit Bay, L. Ont., NY, USA;	42.328,	-82,472,
			TAVgr07-02, TAVgr07-03, TAVgr07-14, TAVgr07-15, TAVgr07-16,		L. Ont., Rochester, NY, USA;	43.200,	-77.526,
			TAVgr07-10, TAVgr07-11,		Sodus Bay, L. Ont., NY, USA;	43.233,	-76.650,
			TAVgr09-32, TAVgr09-33,		Apostle Is., L. Super., WI, USA;	42.257,	-76.966,
			TAVgr07-05,		Baseline L., MI, USA;	47.085,	-90.641,
			TAVgr09-13		Budd L., MI, USA;	42.427,	-83.899,
					Cayuga-Seneca Canal, NY, USA;	44.015,	-84.788,
					Ransomville, NY, USA;	42.910,	-76.910,
					Clear Fork Res., OH, USA;	43.240,	-79.920,
					L. Winnebago, WI, USA;	40.716,	-82.643,
					Oak Ck./Grant Pk., WI, USA	44.028,	-88.421,
Lepomis gibbosus	TAVgr07-04 (vcG006)	HQ623438	__	“ “	Budd L., MI, USA	42.926	-87.770
L. macrochirus	TAVgr07-01 (vcG005)	HQ623437	__	“ “	“ “	44.015	-84.788
“ “	TAVgr07-20 (vcG007)	HQ623439	__	“ “	East Harb., L. Erie, OH, USA	“ “	“ “
Micropterus dolomieu	TAVgr07-24 (vcG008)	HQ623440	__	“ “	Sturgeon Bay, L. Michigan, MI, USA	41.541	-82.789
						44.884	-87.388

		M. salmoides	*LMB	XXXXXX	__	This study	Sandusky B., L. Erie, USA	41.453	-82.726
		Morone chrysops	TAVgr06-16 (vcG004)	HQ623436		" "	Western Basin, L. Erie, OH, USA;	41.492,	-81.667,
					TAVgr08-01		Clear Fork Res., OH, USA	40.716	-82.643
		Neogobius melanostomus	t. Goby 1-5	AB672615	__	Ito et al., (2012)	Cape Vincent, L. Ontario, NY, USA	44.126	-76.334
		" "	o. vcG014	__	__	Cornwell et al., (2012)	Selkirk, St. Lawrence R., USA	43.577	-76.203
		" "	p. TAVgr10-10 (vcG015)	__		" "	Cape Vincent, St. Lawrence R., USA;	44.185,	-76.224,
					TAVgr10-08		Selkirk Shores St. Pk., L. Ontario, USA	43.577	-76.203
		" "	q. vcG016	__	__	" "	Selkirk, St. Lawrence R., USA	" "	" "
		" "	s. GL2010-098	KC117250	__	Garver et al., (2013)	L. Simcoe, Ontario, CAN	44.437	-79.339
		Perca flavescens	TAVgr09-17 (vcG010)	HQ623442	__	" "	L. Michigan, WI, USA	43.039	-87.802
		Percopsis omiscomaycus	m. TAVgr06-53 (vcG011)	HQ623443	__	" "	Central Basin, L. Erie, OH, USA	41.755	-81.286
		" "	n. TAVgr10-02 (vcG013)	__	__	Kurath et al., (unpub.)	Hammond Bay, L. Michigan, MI, USA	45.518	-84.085
	IVc	*Fundulus heteroclitus*	CA-NB00-01	EF079896	__	Gagné et al., (2007)	Ruisseau George Collette, near Bouctouche, NB, CAN	46.450	-64.682
		Gasterosteus aculeatus	CA-NB00-02	HQ168405	__	" "	" "	" "	" "
		Morone saxatilis	CA-NB04-01b	HQ453208	__	" "	Miramichi Bay, Baie du Vin, NB, CAN	47.163	-64.574
		" "	CA-NB02-01	EF079897		" "	" ",	" ",	" ",
					CA-NB04-01 (EF079898)		Miramichi R., NB, CAN	47.004	-65.541
		S. trutta	CA-NS04-01	EF079899	__	" "	French R., NS, CAN	45.576	-62.425
Nv	**I**	*Oncorhynchus kisutch*	Hededam	Z93412	__	Stone et al., (1997)	Spjarup Hededamme, DNK	56.648	9.271
	II	*Ciliata mustela*	*DK-1p49	DQ159193	__	Einer-Jensen et al., (2005)	Baltic S., DEU	56.500	19.000

III	*Salmo trutta*	strain 23-75	FN665788		Biacchesi et al., (2010), Einer-Jensen et al., (2005)	FRA	__	__
				FR-2375 (DQ159196)				
IVa	*Oncorhynchus kisutch*	Makah	U28745		Basurco and Benmansour (1995), Einer-Jensen et al., (2005)	Makah National Fish Hatchery, Neah Bay, WA, USA	48.290	-124.650
				US Makah (DQ159204)				
	" "	JF00Ehi1	AB490792	__	Ito et al., (unpub.)	Ehime, JPN	33.750	132.600
	Paralichthys olivaceus	KRRV9822	AB179621	__	Byon et al., (unpub.)	JPN	__	__
	" "	KJ2008	JF792424	__	Kim and Kim (unpub.)	S. KOR	__	__
	" "	KR-CJA	JQ651389		Lee et al., (unpub.)	Wando, KOR,	34.389,	126.702,
				KR-YGH (JQ651394)		Jeju, KOR	33.430	126.546
IVb	*Ambloplites rupestris*	*TAVgr08-03[a] (vcG009)	XXXXXX	__	This study	N. Point Marina, L. MI, IL, USA	42.499	-90.689
	Aplodinotus grunniens	*DRUM	XXXXXX	__	" "	Sandusky B., L. Erie, USA	41.453	-82.726
	Cyprinus carpio	*TAVgr07-12[a] (vcG003)	XXXXXX	__	" "	Dunkirk Harb., L. Erie, NY, USA	42.490	-79.338
	Esox masquinongy	MI03GL	GQ385941		Elsayed et al., (2006), Ito et al., (2012)	L. St. Clair, MI, USA;	42.391,	-82.911,
				Goby 1-5 (AB672615)		Cape Vincent, L. Ontario, NY, USA	44.126	-76.334
	Morone chrysops	*TAvgr06-16[a] (vcG004)	XXXXXX	__	This study	Western B., L. Erie, USA	41.492	-81.667
	Micropterus dolomieu	*TAVgr07-24[a] (vcG008)	XXXXXX	__	" "	Sturgeon Bay, L. Michigan, MI, USA	44.884	-87.388
	M. salmoides	*LMB	XXXXXX	__	" "	Sandusky B., L. Erie, USA	41.453	-82.726
	Lepomis gibbosus	*TAVgr07-04[a] (vcG006)	XXXXXX	__	" "	Budd L., MI, USA	44.015	-84.788
	L. macrochirus	*TAVgr07-01[a] (vcG005)	XXXXXX	__	" "	" "	" "	" "
	" "	*TAVgr07-20[a] (vcG007)	XXXXXX	__	" "	East Harb., L. Erie, OH, USA	41.541	-82.789
	Perca flavescens	*TAVgr08-02[a] (vcG002)	XXXXXX	__	" "	L. Michigan, near Milwaukee, USA	42.799	-87.761
	" "	*TAVgr09-17[a] (vcG010)	XXXXXX	__	" "	L. Michigan, WI, USA	43.039	-87.802

	IVc	*Gasterosteus aculeatus*	*CA-NB00-02[c]	XXXXXX	__	“ “	Ruisseau George Collette, near Bouctouche, NB, CAN	46.450	-64.682
P	**I**	*Oncorhynchus mykiss*	Hededam	Z93412	__	Stone et al., (1997)	Spjarup Hededamme, DNK	56.648	9.271
	II	*Clupea harengus*	*FI-ka663-03[b]	XXXXXX	__	This study	Archipelago S., FIN	60.290	21.290
	III	*Salmo trutta*	strain 23-75	FN665788	__	Biacchesi et al., (2010)	FRA	__	__
	IVa	*Onchorhychus kisutch*	Makah	U02630	__	Benmansour et al., (1994)	Makah National Fish Hatchery, Neah Bay, WA, USA	48.290	-124.650
		Paralichthys olivaceus	KJ2008	JF792424	__	Kim and Kim (unpub.)	S. KOR	__	__
		“ ”	Isolate #25	AB060726	__	Nishizawa et al., (unpub.)	Wakasa Bay, JPN	35.200	134.240
		“ ”	KRRV9822	AB179621	__	Byon et al., (unpub.)	JPN	__	__
	IVb	*Esox masquinongy*	MI03GL	GQ385941	*DRUM(XXXXX), *TAVgr07-12[a]; vcG003(XXXXX), *TAVgr07-01[a]; vcG005(XXXXX), *TAVgr07-04[a]; vcG006(XXXXX), *TAVgr07-20[a]; vcG007(XXXXX), *TAVgr07-24[a] vcG008(XXXXX), *TAVgr08-03[a]; vcG009(XXXXX) *TAVgr09-17[a] vcG010(XXXXX),	Elsayed et al., (2006), This study	L. St. Clair, MI, USA; Sandusky B., L. Erie, USA; Dunkirk Harb., L. Erie, NY, USA; Budd L., MI, USA; East Harb., L. Erie, OH, USA; Sturgeon Bay, L. Michigan, MI, USA; N. Point Marina, L. MI, USA; L. Michigan, WI, USA	42.391, 41.453, 42.490, 44.015, 41.541, 44.884, 42.499, 43.039	-82.911, -82.726, -79.338, -84.788, -82.789, -87.388, -90.689, -87.802
		Micropterus salmoides	*LMB	XXXXXX	__	“ “	Sandusky B., L. Erie, USA	41.453	-82.726
		Morone	*TAvgr06-16[a]	XXXXXX	__	“ “	Western B., L. Erie,	41.492	-82.667

		chrysops	(vcG004)				USA		
		Neogobius melanostomus	t. Goby1-5	AB672615	__	Ito et al., (2012)	St. Lawrence R., near Cape Vincent, NY, USA	44.383	-75.868
		Perca flavescens	*TAVgr08-02[a] (vcG002)	XXXXXX	__	This study	L. Michigan, near Milwaukee, USA	42.799	-87.761
	IVc	*Gasterosteus aculeatus*	*CA-NB00-02[c]	XXXXXX	__	" "	Ruisseau George Collette, near Bouctouche, NB, CAN	46.450	-64.682
M	**I**	*Onchorhychus mykiss*	Hededam	Z93412	__	Stone et al., (1997)	Spjarup Hededamme, DNK	56.648	9.271
	III	*Salmo trutta*	strain 23-75	FN665788	__	Biacchesi et al., (2010)	FRA	__	__
	IVa	*Onchorhychus kisutch*	Makah	U03503	__	Benmansour et al., (1994)	Makah National Fish Hatchery, Neah Bay, WA, USA	48.290	-124.650
		Paralichthys olivaceus	KJ2008	JF792424	__	Kim and Kim (unpub.)	S. KOR	__	__
		" "	JF00Ehi1	AB490792	__	Ito et al., (unpub.)	Ehime Prefecture, JPN	33.852	132.770
		" "	KRRV9822	AB179621	__	Byon et al., (unpub.)	JPN	__	__
		" "	KR-CJA	JQ651387	KR-YGH (JQ651392)	Lee et al., (unpub.)	S. KOR	__	__
	IVb	*Aplodinotus grunniens*	*DRUM	XXXXXX	__	This study	Sandusky Bay, L. Erie, USA	41.453	-82.726
		Esox masquinongy	MI03GL	GQ385941	Goby 1-5 (AB672615),	Elsayed et al., (2006), Ito et al., (2012),	L. St. Clair, MI, USA;	42.391,	-82.911,
							Cape Vincent, L. Ontario, NY, USA;	44.126	-76.334
					*TAVgr07-12[a]; vcG003 (XXXXX),	This study	Dunkirk Harb., L. Erie, NY, USA	42.490,	-79.338,
					*TAvgr06-16[a]; vcG004 (XXXXX),		Western B., L. Erie, USA;	41.492,	-81.667,
					*TAVgr07-01[a]; vcG005 (XXXXX),		Budd L., MI, USA;	44.015,	-84.788,
					TAVgr07-04[a]; vcG006 (XXXXX),				

				*TAVgr07-20[a]; vcG007 (XXXXX), *TAVgr08-03[a]; vcG009 (XXXXX)			East Harb., L. Erie, OH, USA; North Point Marina, L. MI, USA	41.541, 42.499	-82.789, -90.689
	Micropterus dolomieu	*TAVgr07-24[a] (vcG008)	XXXXXX	__	"	"	Sturgeon Bay, L. Mich., MI, USA	44.884	-87.388
	M. salmoides	*LMB	XXXXXX	__	"	"	Sandusky Bay, L. Erie, USA	41.453	-82.726
	Perca flavescens	*TAVgr08-02[a] (vcG002)	XXXXXX	__	"	"	L. Mich., near Milwaukee, USA	42.799	-87.761
	" "	*TAVgr09-17[a] (vcG010)	XXXXXX	__	"	"	L. Mich., WI, USA	43.039	-87.802
IVc	*Gasterosteus aculeatus*	*CA-NB00-02[c]	XXXXXX	__	"	"	Ruisseau George Collette, near Bouctouche, NB, CAN	46.450	-64.682

Chapter 4

A new StaRT-PCR approach to detect and quantify fish Viral Hemorrhagic Septicemia virus (VHSv): Enhanced quality control with internal standards

Previously published as Pierce, L.R., Willey, J.C., Crawford, E.L., Palsule, V.V., Leaman, D.W., Faisal, M., Kim, R.K., Shepherd, B.S., Stanoszek, L.M., and Stepien, C.A. (2013) A new StaRT-PCR approach to detect and quantify fish Viral Hemorrhagic Septicemia virus (VHSv): Enhanced quality control with internal standards. *Journal of Virological Methods*

4.1 ABSTRACT: Viral Hemorrhagic Septicemia virus (VHSv) causes one of the world's most important finfish diseases, killing >80 species across Eurasia and North America. A new and especially virulent strain (IVb) emerged in the North American Great Lakes in 2003, threatening fisheries, baitfish, and aquaculture industries. Weeks-long and costly cell culture is the OIE and USDA-APHIS approved diagnostic. A new Standardized Reverse Transcriptase Polymerase Chain Reaction (StaRT-PCR) assay that uniquely incorporates internal standards to improve accuracy and prevent false negatives was developed and evaluated for its ability to detect and quantify VHSv. Results from StaRT-PCR, SYBR® green real time qRT-PCR, and cell culture were compared, as well as the effects of potential PCR inhibitors (EDTA and high RNA). Findings show that StaRT-

PCR is sensitive, detecting a single molecule, with 100% accuracy at six molecules, and had no false negatives. In comparison, false negatives ranged from 14-47% in SYBR® green real time qRT-PCR tests, and 47-70% with cell culture. StaRT-PCR uniquely controlled for EDTA and RNA interference. Range of VHSv quantitation by StaRT-PCR was $1.0x10^0$-$1.2x10^5$ VHSv/10^6 *actb1* molecules in wild caught fishes and $1.0x10^0$-$8.4x10^5$ molecules in laboratory challenged specimens. In the latter experiments, muskellunge with skin lesions had significantly more viral molecules (mean=$1.9x10^4$) than those without ($1.1x10^3$) ($p<0.04$). VHSv infection was detected earlier in injection than in immersion challenged yellow perch (two versus three days), with molecule numbers in both being comparable and relatively consistent over the remaining course of the experiment. Our results show that the StaRT-PCR test accurately and reliably detects and quantifies VHSv.

4.2 INTRODUCTION

4.2.1 VHSv characteristics and spread

Viral Hemorrhagic Septicemia virus (VHSv) is one of the world's most serious fish pathogens, killing over 80 marine and freshwater species (including trout, salmon, and perch) across the Northern Hemisphere (Faisal et al., 2012), yet has lacked a rapid and accurate diagnostic test. The virus is a negative-sense, single stranded RNA *Novirhabdovirus* of ~12,000 nucleotides, with six open reading frames of 3'*N-P-M-G-Nv-L*'5 (Ammayappan and Vakharia, 2009). A new and especially virulent VHSv

substrain (IVb) emerged in the Laurentian Great Lakes of North America in 2003, causing massive fish kills (Elsayed et al., 2006; Faisal et al., 2012) that have threatened the fisheries, aquaculture, baitfish, and tourism industries (Leighton, 2011).

Transmission of VHSv occurs via fish waste, reproductive fluids, and skin secretions. Its viral particles can live up to 13 days in the water (Hawley and Garver, 2008) and are transported via boating, ballast water, fishing tackle, and animals – e.g., amphipod crustaceans, leeches, turtles, and birds (Faisal and Schulz, 2009; Bain et al., 2010; Faisal and Winters, 2011; Goodwin and Merry, 2011). Clinical signs of infection vary, ranging from erratic swimming, exophthalmia (bulging eyes), distended abdomens, to extensive external/internal bleeding (Winton and Einer-Jensen, 2002). Since November 2006, the eight U.S. states (Illinois, Indiana, Ohio, Pennsylvania, Michigan, Minnesota, New York, Wisconsin) and two Canadian provinces (Ontario and Quebec) that surround the Great Lakes have required that fish are certified as VHSv-free prior to interstate transport (Aquatic Invasive Species Action Plan, 2011), for which the Office International des Epizooties (OIE) recommends a month long cell culture process (OIE, 2009).

VHSv (originally called "*Nierenschwellung*") first was described from European salmonid aquaculture (Schäperclaus, 1938), and was isolated in 1962 from infected rainbow trout (*Oncorhynchus mykiss*) (Einer-Jensen et al., 2004). Four genetically distinct strains (I-IV) and various substrains have been recognized (Snow et al., 1999; Einer-Jensen et al., 2004), whose phylogenetic and biogeographic relationships recently were analyzed by Pierce and Stepien (2012). Strains I-III are found in Europe, where strain I infects >13 freshwater species including rainbow and brown trout (*Salmo trutta*).

Strain II comprises a tight genetic cluster (Pierce and Stepien, 2012), which primarily infects Pacific herring (*Clupea pallasii*) in the Baltic Sea and Finland Archipelago (Gadd et al., 2011). Strain III is distributed peripherally to strain I, infecting Atlantic cod (*Gadus morhua*), European eel (*Anguilla anguilla*), haddock (*Melanogrammus aeglefinus*), Norway pout (*Trisopterus esmarki*), Pacific herring, rainbow and brown trout, turbot (*Scophthalmus maximus*), whiting (*Merlangius merlangus*), and others. Strain IV was described in 1988 from the North American Pacific Northwest (substrain a), where it is found in a wide variety of marine fish species (including chinook salmon (*Oncorhynchus tshawytsha*), coho salmon (*Oncorhynchus kisutch*), Pacific cod (*Gadus macrocephalus*), Pacific sardine (*Sardinops sagax*), and smelt (*Thaleichthys pacificus*)), with a few occurrences in Japan and Korea (in black seabream (*Spondyliosoma cantharus*) and olive flounder (*Paralichthys olivaceus*)) (Kim and Faisal, 2011). Substrain IVb was identified in the Great Lakes (Lake St. Clair) basin from an adult muskellunge (*Esox masquinongy*) collected in 2003 (MI03GL; Elsayed et al., 2006). IVb has spread to infect 31 freshwater fish species across all five of the Great Lakes (Thompson et al., 2011), with some isolates found in invertebrates, indicating that they might serve as transmission vectors (Faisal and Schulz, 2009; Faisal and Winters, 2011).

4.2.2 Need for a new VHSv diagnostic test

Screening methods for VHSv that are approved by the World Organization for Animal Health (OIE, 2009), and the Fish Health Section of the U.S. Fish and Wildlife Service and the American Fisheries Society (2010) are based on identification via cell

culture, followed by confirmation either with reverse transcriptase PCR or serologically via an indirect fluorescent antibody test or an enzyme-linked immunosorbent assay. Although those methods readily detect high concentrations of the virus, they may fail to detect low levels of virus in carrier fish. Notably, fish have been shown to shed virus for up to 15 weeks post infection (Kim and Faisal, 2012).

Cell culture identification and subsequent confirmation is labor intensive, time consuming (up to four weeks), and less sensitive than direct PCR-based detection, with false negative rates reported up to 95% (López-Vázquez et al., 2006; Miller et al., 1998; Winton and Einer-Jensen, 2002). For fish farms, hatcheries, and baitfish operators, a month-long holding period for viral detection leads to economic loss and possible viral spread. Additionally, cell culture facilities typically "pool" samples (i.e., tissues from several fish samples are homogenized together), which dilutes the number of viral particles and circumvents pinpointing exactly which samples are VHS positive. Further, cell culture identification methods, such as plaque assay, may not precisely quantify the number of viral particles or allow determination of the amount leading to infection.

Non-lethal techniques to detect neutralizing glycoprotein (*G*) gene antibodies in fish blood have been developed, with experiments showing that VHSv remained detectable for ≥90 days post exposure (Millard and Faisal, 2012). However, that antibody approach relied on confluent cultured cells followed by a 6-7 day incubation and an optimum virus concentration for stimulating antibody production (which remains to be characterized). Thus, that weeks-long test may be insufficient for detecting low amounts of VHSv in carrier fish. An assay using monoclonal antibodies by Ito et al. (2012) was developed to identify VHSv strains and substrains. However, that method also used

cultured cells and relied on generation of specific antibodies, which are lengthy and costly procedures.

Quantitative real-time PCR (qRT-PCR) approaches have greater sensitivity and reduced detection time in comparison to cell culture and plaque assays (Bruchhof et al., 1995; Miller et al., 1998; Guillou et al., 1999). Several qRT-PCR assays for VHSv have been developed, including: Chico et al. (2006), López-Vázquez et al. (2006), Liu et al. (2008), Matejusova et al. (2008), Cutrín et al. (2009), Hope et al. (2010), Garver et al. (2011), Jonstrup et al. (2012), and Phelps et al. (2012); however, all of these have accuracy limitations. Just three of those tests quantified VHSv levels: Liu et al. (2008), Hope et al. (2010), and Garver et al. (2011). For example, Liu et al. (2008) distinguished to 140 viral copies of a single VHSv type (Ia; isolate Fil3); however, it remains unknown whether their test method, which was based on the *G*-gene, would work on other VHSv variants. Other assays used the nucleoprotein (*N*) gene (Chico et al., 2006; Cutrín et al., 2009; Garver et al., 2011; Jonstrup et al., 2012; López-Vázquez et al., 2006; Matejusova et al., 2008; OIE 2009), including a test by the Hope et al. (2010) that reliably distinguished to 100 viral copies. However, higher amounts of RNA (1 μg) led to a 20-fold reduction in their PCR amplification signal, with further signal decline at RNA concentrations of 4-8 μg (Hope et al., 2010). That decline likely was due to interference and/or reagent carry over and would be prone to generate false negative results. The assay by Garver et al. (2011) detected to 100 viral copies, based on results from several laboratories in blind experiments. Jonstrup et al. (2012) purported increased sensitivity compared to other assays because they observed amplification at a lower cycle threshold (C_t; the number of cycles at which the fluorescence exceeded the threshold). Their

method was evaluated for all four VHSv strains and 79 isolates, yet did not discern significantly more VHS positives than cell culture (χ^2=0.10, df=1, NS; statistical analyses performed in the present study using their results). Their findings suggested a large proportion of false negatives (Jonstrup et al., 2012).

4.2.3 The StaRT-PCR method and study objectives

To date, qRT-PCR and plaque/immunological assays for VHSv have lacked Internal Standards (IS) to control for interfering substances and false negative results, which may lead to misdiagnosis and potential viral spread. The present study thus developed a new PCR-based test for VHSv, which incorporates a standardized mixture of internal standards (SMIS) in a Standardized Reverse Transcriptase Polymerase Chain Reaction (StaRT-PCR); this method follows the StaRT-PCR approach outlined by Crawford et al. (2002), Willey et al. (2004), and Canales et al. (2006).

StaRT-PCR is a form of competitive PCR (Gilliland et al., 1990; Celi et al., 1993) that measures expression of each gene relative to a known number of copies of a synthetic competitive template IS within a SMIS (Willey, 2004). An IS is constructed for each target gene (i.e., VHSv *N*-gene) and for one or more reference genes (i.e., genes that are expressed at a relatively consistent level across many tissues and conditions), which are combined into a SMIS. The SMIS is loaded into each reaction (rxn). This approach also controls for interfering substances, such as PCR inhibitors, and prevents false negative results. For example, if the IS PCR product is not observed, it will be interpreted as a failed test and not as valid assessment for the absence of VHSv. StaRT-PCR is

designed to yield rapid, reproducible, standardized, and quantified measures from several genes simultaneously.

The research aim of the present study was to evaluate the performance and accuracy of the newly developed StaRT-PCR VHSv test that incorporates synthetic IS, in comparison to conventional qRT-PCR-based assays (e.g., SYBR® green qRT-PCR) that lack IS, and cell culture. Experiments were conducted to test the ability of StaRT-PCR to discern and quantify VHSv from (A) pellets of cells infected *in vitro*, (B) wild caught fishes, (C) fish (muskellunge and yellow perch (*Perca flavescens*)) that were experimentally challenged with VHSv, and (D) laboratory altered samples containing various PCR inhibitors. The latter included both exogenous inhibitors (e.g. Ethylenediaminetetraacetic acid (EDTA)) and endogenous inhibitors from high levels of RNA in reverse transcription. The experiments were designed to discern if StaRT-PCR is (1) VHSv-specific, (2) accurately measures the amount of VHSv in infected fish cells, (3) controls for PCR inhibitors, (4) accurately and reliably diagnoses and quantifies VHSv in field and laboratory fish samples, and (5) has greater sensitivity and reliability than cell culture or other qRT-PCR methods. Laboratory challenge experiments also were conducted to determine whether the number of VHSv molecules significantly differs in challenged fish (6) with or without clinical signs of infection, (7) over the course of early infection (to day six), and/or (8) between immersion and injected challenged individuals. These evaluations provided examples of the potential applications of the StaRT-PCR VHSv test.

4.3 MATERIALS AND METHODS

4.3.1 Design of the StaRT-PCR test

Primers and the IS designed for the StaRT-PCR VHSv test targeted the central portion of the VHSv *N*-gene (Fig. 4-1). The *N*-gene expresses RNA transcripts most abundantly in rhabdovirus infections (Chico et al., 2006) and is relatively conserved across species due to its RNA-binding function (Hope et al., 2010). Other studies showed that primers targeting the VHSv *N*-gene were more efficient and sensitive than those targeting the *G*-gene (Chico et al., 2006; Cutrín et al., 2009). In the present study, *N*-gene sequences were aligned from all VHSv variants, related Novirhabdoviruses, and other viruses from NIH GenBank (http://www.ncbi.nlm.nih.gov/genbank/) using the BLAST procedure (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to ensure that the designed primers did not recognize other viruses. VHSv-specific inter-variant homologous sequences then were targeted for primer design, and Oligo software (http://www.oligo.net/) was used to select primers based on the absence of stable duplex formation, low likelihood of false priming sites, and an optimal annealing temperature of 58^{o}C. Primer sequences are given in Table 4.1, and their locations and relationships to those from other assays are shown in Fig. 4-1.

Primer sets and IS for three fish reference genes – ß-actin (*actb1*), elongation factor 1 alpha subunit (*ef1a*), and 18S ribosomal RNA (*18srRNA*) – were developed to ensure accurate diagnosis and quantitation under a variety of PCR and sample conditions (Table 4.1). Reference gene sequences for primers and IS were selected from conserved

regions across 10 VHSv-affected fish species (five sport fish species: yellow perch, smallmouth bass (*Micropterus dolomieu*), round goby (*Neogobius melanostomus*), freshwater drum (*Aplodinotus grunniens*), and walleye (*Sander vitreus*) and five baitfish species: emerald shiner (*Notropis atherinoides*), golden shiner (*Notemigonus crysoleucas*), fathead minnow (*Pimephales promelas*), spottail shiner (*Notropis hudsonius*), and alewife (*Alosa pseudoharengus*)). Each reference gene was measured to determine whether it was expressed at a consistent level in VHSv infected and non-infected fish, to validate its use as a control for variation in loading of sample into PCR. In these experiments, the reference genes were measured for five VHS positive and five negative control fish samples (muskellunge and yellow perch).

The IS for the VHSv *N*-gene and each of the three reference genes were prepared following the method of Celi et al. (1993; see their Fig. 1a). First, primers were designed and synthesized for each of the four genes. Then, a modified reverse primer, called the IS primer, was constructed such that the 3' end had sequence homology to a region between the forward and reverse primers and the 5'end of the IS primer was identical to the reverse primer (Table 4.1). PCR extension only occurred from the sequence annealed at the 3' end. This resulted in a PCR product for the IS that was shorter than the VHSv native target (NT) sequence. This difference in size enabled electrophoretic size separation and quantitation of IS and NT. The IS primer was used only to create the IS and was not used in the StaRT-PCR assays.

To develop the SMIS, each IS was generated by separate PCR amplification of fish cDNA using the forward primer (19-21 bp; Table 4.1) and the IS primer (38-42 bp; Table 4.1) in five 10 μl PCR replicates containing 1 μl (0.05 μg) of each primer, 0.5 U/μl

Go-TAQ polymerase (Promega, Madison, WI), 1 μl 10X $MgCl_2$ PCR buffer, 1 μl 0.2 mM dNTPs, 5.5 μl RNase-free water, and 1 μl cDNA resulting from a 90 μl reverse transcription rxn containing 1 μg of VHSv positive fish RNA. Each of the five replicate samples then were transferred to LightCycler® Capillaries (Roche, Indianapolis, IN) and cycled on a Rapid Cycler 2 (Idaho Technology, Inc., Salt Lake City, UT), with each cycle being 5 sec at 94^{o}C, 10 sec at 58^{o}C, and 15 sec at 72^{o}C, and a slope of 9.9 for rapid temperature change between cycles (35 cycles total). PCR products from the five replicates then were combined, separated by electrophoresis on a 2% low melting agarose gel (Fisher Scientific, Fair Lawn, NJ) containing 5 μl ethidium bromide (10 mg/ml) per 100 ml 1X TRISAcetate EDTA buffer, and visualized on a UV transilluminator. The band corresponding to each IS was excised from the gel and purified using a Qiaquick Gel Extraction Qiagen Kit (Qiagen, Germantown, MD). Molarities were calculated by analysis of 1 μl of the purified products on a Agilent 2100 Bioanalyzer (Agilent Technologies, Inc, Santa Clara, CA) and concentrations were converted into molecules using the below formula (1):

(1) [Moles/Liter]/[$1.0x10^{6}$ μl/L]x[$6.0x10^{23}$ molecules/M]=molecules/μl

A Standardized Mixture of Internal Standards (SMIS) was created by mixing the individual gene IS together, as shown in Table 4.2, to enable measurement of expression across a range of possible numbers of target molecules. To obtain SMIS "A", a stock was created such that the IS for *18srRNA* was at 10^{-10} M, the IS for *actb1* was at 10^{-11} M, the IS for *ef1a* was at 10^{-10} M, and the IS for VHSv was at 10^{-10} M. A separate mixture of the

three reference gene IS, which did not contain VHSv IS, was created with *18srRNA* at 10^{-10} M, *actb1* at 10^{-11} M, and *ef1a* at 10^{-10}. This reference gene IS mixture was used in a 10-fold serial dilution of the SMIS "A" stock to create stocks for SMIS B-H (row 1 of Table 4.2). A solution of 0.1 ng/µl yeast tRNA carrier (Invitrogen, Carlsbad, CA) was used in all dilutions to prevent adherence of the negatively charged IS molecules to the plastic tube or pipette tip surfaces/diluents. Stock SMIS then was diluted down serially with carrier solution (rows 2-8 of Table 4.2) to generate a range of concentrations and enable measurement over several orders of magnitude.

4.3.2 How to perform StaRT-PCR

StaRT-PCR methods were used as described by Willey et al. (1998, 2004). First, 0.25-0.50 g fish tissue (spleen was preferred since it is readily identifiable by collectors unfamiliar with fish anatomy) was ground under liquid nitrogen using a sterile mortar and pestle, and RNA was extracted using TriREAGENT® (Molecular Research Center, Inc., Cincinnati, OH) following the manufacturer's protocol. Second, RNA was re-suspended in 30 µl RNase-free water, quantified with a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA), and the concentration was adjusted to 1 µg RNA/µl. A 30 µl RNA volume was treated with DNA-free DNase and DNAse Removal Reagents (Ambion Life Technologies, Grand Island, NY) to remove any contaminating gDNA. Third, reverse transcription of this purified RNA to cDNA was conducted using 1 µg RNA, 5X First Strand buffer, 10 mM dNTPs, 0.05 mM random hexamers, 25 U/µl RNasin, and 200 U/µl M-MLV in a 90 µl rxn volume. Reverse transcription rxns were

carried out at 94°C for 5 min, 37°C for 1 h, and 94°C for 5 min. All cDNA was stored at -20°C until further use.

For each sample assessed, an initial set of PCR amplifications was conducted to determine how much cDNA and how much reference gene (*actb1*) IS should be combined to achieve approximately (within a 10-fold range) a 1:1 ratio between the NT and IS for *actb1*. For subsequent reactions, the *actb1* NT and IS concentrations were held constant for each sample, while the SMIS used (A-H) varied (Table 4.2) to achieve approximately (within a 10-fold range) a 1:1 ratio of the VHSv NT relative to the VHSv IS. For example, if the SMIS "D" concentration was used to assay an unknown quantity of VHSv from a fish sample, and the NT amount of VHSv was found to be >10-fold more than the VHSv IS, the next step would be to repeat the experiment using the same amount of cDNA and substituting with SMIS "C", which has a 10-fold higher concentration of VHSv IS (Willey et al., 2004).

StaRT-PCR was conducted in 10 µl rxn volumes as follows:

(A) Primer pairs for each gene were mixed together (each at 0.05 µg/µl) to decrease pipetting error, then 1 µl of the mixture for the VHSv target gene and *actb1* reference gene was placed in separate tubes.

(B) A master mixture was prepared containing both the appropriate concentration of cDNA and the appropriate SMIS (as determined in the initial set of PCR) to ensure equal loading of both into the separate rxn tubes for measurement of VHSv or *actb1*. Also included in the master mix were: Go-TAQ polymerase (final conc. of 0.1 U/µl), 10X PCR buffer containing 30 mM $MgCl_2$ (final conc. of 1X with 3 mM $MgCl_2$), dNTPs (final conc. of 0.2 mM), and RNase-free water.

(C) 9 μl of the master mixture was added into each of the tubes from (A), mixed, transferred into LightCycler® capillaries, and then PCR-amplified on a Rapid Cycler 2, as described in section 2.1.

(D) To check for contamination in the PCR reagents and to confirm that reactions worked, a total of seven additional tubes were prepared for each experiment: (1) SMIS only (no cDNA) with primers for each gene, (2) a known VHSv positive and negative fish with primers for each gene, and (3) a water/reagent control with primers for each gene (nuclease-free H_2O).

(E) PCR products were visualized on an Agilent 2100 Bioanalyzer, as above.

(F) This process (A-E) was repeated 3X to determine a mean and standard deviation, and calculate the standard error, for the number of VHSv molecules/10^6 *actb1* using the equations below (2 and 3).

(2) Correcting NT product size (this is necessary because quantitation was by optical density of intercalator dye and this, in turn, was related to molecule length as well as copy number).

(a) [Expected IS bp/Expected NT bp]x[NT Area Under Curve on Agilent graph]=Corrected NT Area Under Curve

(b) [Corrected NT Area Under Curve/IS Area Under Curve]x[number of IS molecules in the rxn]=number of NT molecules

(3) Normalizing VHSv to *actb1*

[Number of NT molecules VHSv/Number of NT molecules *actb1*]x[10^6]=VHSv molecules/10^6 *actb1* molecules

4.3.3 Specificity, linearity, precision, and accuracy of the VHSv StaRT-PCR test

The VHSv StaRT-PCR assay was tested for non-specific amplification of several viruses (Table 4.3) in separate experiments. Viruses assessed included the human Encephalomyocarditis virus and Vesicular Stomatitis virus, and five fish viruses related to VHSv – Hirame Rhabdovirus, Infectious Hematopoietic Necrosis virus, Infectious Pancreatic Necrosis virus, Spring Viremia of Carp virus, and Snakehead Rhabdovirus. The latter is sister species of VHSv and has 62% similarity (Ammayappan and Vakharia, 2009; Pierce and Stepien, 2012). Additionally, all four strains of VHSv were tested (I-IV), including 25 VHSv isolates and European, Asian, and North American variants (Table 4.3). Viruses were obtained either as cell culture supernatant, RNA, or tissue from infected fishes. These samples were processed following the same procedure, except that cell culture supernatants were extracted using TriREAGENT-LS® (Molecular Research Center Inc.). All samples were assayed in triplicate.

Linearity of the StaRT-PCR assay was tested by inoculating the *Epithelioma papulosum cyprini* (EPC) cell line with known amounts of VHSv. The EPC cell line originated from carp (*Cyprinus carpio*) and was obtained from the American Type Culture Collection (ATCC; Manassas, VA). Cells were propagated following Kim and

Faisal (2010), trypsinized from the plate once deemed confluent, and collected as pellets after centrifugation. Pellets (three per dilution) then were spiked with a known dilution of 10^0-10^5 pfu VHSv-IVb (strain MI03GL)/10^6 cells. Two cell pellets containing nuclease-free H_2O served as negative controls. StaRT-PCR was used to measure VHSv in each sample in triplicate (totaling nine measurements per dilution). A log-log (i.e., power) regression analysis was employed to determine whether the number of molecules detected with StaRT-PCR followed a linear trend with viral dilution. Relative accuracy was quantified as the distribution of the percent difference between the number of molecules measured versus the number expected across all dilutions (per Shabir, 2003). Precision, the measure of the degree of repeatability of an analytical method (Shabir, 2003), was assessed by calculating the coefficient of variation for each sample across all three experiments. Values are reported as percentages.

The true accuracy of the StaRT-PCR VHSv assay, defined as the agreement between a measurement and its known value (Shabir, 2003), was evaluated by Poisson distribution analysis (Vogelstein and Kinzler, 1999). According to this method, the laws of chance governing stochastic sampling variation were used to calculate the concentration based on the relationship between the fraction of PCRs observed to be positive relative to the fraction expected. The observations were based on nine separate PCRs each of 16 extreme limiting dilutions of a VHS-IVb sample. VHSv-IVb was prepared from a smallmouth bass isolate MI03GL (a homogenate of spleen, kidney, and brain tissue; provided by P. Bowser, Cornell University College of Veterinary Medicine, Ithaca, NY). MI03GL was the original IVb isolate (Elsayed et al., 2006) and is the most widespread in the Great Lakes (Thompson et al., 2011). PCRs were run on dilutions

expected to contain 0.5, 1, 2, 3, 4, 5, 6, 8, 10, 25, 50, $1.0x10^2$, $2.0x10^2$, $6.0x10^2$, $6.0x10^3$, and $6.0x10^6$ VHSv molecules, which were mixed with the appropriate SMIS to achieve an approximate VHSv NT:IS ratio of 1:1. An exponential regression analysis (in SPSS v21; http://www01.ibm.com/software/analytics/spss/; SPSS Inc., Chicago, IL; Norusis, 2008) was used to calculate the concentration of VHSv in the SMIS based on the observed: expected values. The concentration of the SMIS based on StaRT-PCR limiting dilution Poisson distribution analysis and the concentration measured by NanoDrop Spectrophotometer 2000 analysis of the undiluted stock concentration were statistically compared using a χ^2 test (in Microsoft Excel).

4.3.4 Effect of interfering substances on StaRT-PCR

To evaluate accuracy and performance of StaRT-PCR when subjected to possible interfering substances at the PCR level, RNA from a VHSv-IVb (MI03GL) positive smallmouth bass (also used in 2.3) was treated with DNA-free™ (Ambion Life Technologies, Carlsbad, CA) to remove any contaminating gDNA, and was reverse transcribed to cDNA using 1 μg RNA/90 μl rxn. The cDNA then was spiked with 0.1, 0.5, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, and 2.0 mM concentrations of EDTA per PCR and amplified in three separate StaRT-PCR runs, one series with a SMIS and the other lacking a SMIS, to assess each for the occurrence of false negatives. Products of PCR at each EDTA concentration were measured in two ways. First, the mean number of VHSv-IVb molecules/10^6 *actb1* molecules was calculated for each sample (as detailed in 2.2. "*How to Perform StaRT-PCR*", equations 2 and 3). The number of VHSv molecules

measured at 0 mM EDTA concentration was set as the 100% baseline, serving as the control. A one-way ANOVA (Sokal and Rohlf, 1995) in SPSS was used to test for significant difference in the mean number of viral molecules/10^6 *actb1* measured. Second, the total fluorescent light, measured in fluorescent light units, emitted from the Agilent intercalating dye bound to the two PCR product peaks (NT and IS) was measured on an Agilent 2100 Bioanalyzer and combined. The mean fluorescence values for each sample were calibrated as the percent change from the baseline value (0 mM EDTA control).

Reverse transcription efficiency, which is the fraction of mRNA molecules converted into corresponding cDNA molecules (Bustin and Nolan, 2004; Ståhlberget al., 2004), was analyzed by comparing results from 1 (control), 5, 10, 20, and 30 µg of VHSv RNA/90 µl rxn. The same gDNA-free VHSv-IVb positive sample of RNA used for the EDTA test (above) was used in these experiments. A Reverse Transcription Standards Mixture (RTSM) was constructed using *in vitro*-transcribed RNA standards developed by the External RNA Control Consortium (ERCC; http://www.nist.gov/mml/), termed ERCC 113 and 171, and obtained from M. Salit, National Institutes of Standards and Technology, Gaithersburg, MD (Table 4.1). The concentration of each RNA standard was measured using a NanoDrop 2000 spectrophotometer. The ERCC 113 standard was reverse transcribed into cDNA and the cDNA was quantified using the Agilent 2100 BioAnalyzer. The ERCC 171 standard then was diluted to $1.0x10^{-10}$ M using 100 ng/µl yeast tRNA as a carrier, and combined with the ERCC 113 cDNA to yield a mixture of $4.8x10^3 \pm 1.6x10^2$ ERCC 171 RNA molecules and $1.7x10^3 \pm 1.8x10^2$ ERCC 113 cDNA molecules per 2 µl aliquot. A 2 µl aliquot of the RTSM was included in each of the

reverse transcription rxns. The ERCC 171 RNA was converted into cDNA along with other RNA species in the reverse transcription rxn while the ERCC 113 cDNA remained unaltered. Thus, by comparing the ratio of ERCC 171 and ERCC 113 cDNAs after reverse transcription, using StaRT-PCR, it was possible to calculate the reverse transcription efficiency. The numbers of VHSv, *actb1*, and *ef1a* molecules also were quantified in all samples using equations 2 and 3 from "*How to Perform StaRT-PCR*" (above). One-way ANOVAs and *t*-tests (Sokal and Rohlf, 1995) were used to identify possible differences in reverse transcription efficiency and variations in the numbers of measured molecules among the five concentrations.

4.3.5 Laboratory challenged fish experiments

VHSv laboratory challenge experiments were conducted at the Fish Health Laboratory Containment Facility at Michigan State University (MSU). Certified VHSv-free juvenile muskellunge were obtained from the Rathburn National Fish Hatchery (Moravia, Iowa) and maintained in covered 1,900 L tanks in UV-sterilized 12±1°C oxygenated water under controlled ambient light for three weeks acclimation. Muskellunge were fed VHSv-free fathead minnows (Robinson Wholesale, Inc.) that had been determined to be VHS-free using OIE (2009) approved tissue culture isolation assays with confirmation qRT-PCR.

A subset of the 360 muskellunge (20.0±10.9 g; total length 17.1±1.5 cm) were randomly selected and assigned into four test groups (90 fish per group). Each group was challenged via water immersion for 90 min in 37.8 L with either: (1) a low dose of 100

pfu/ml VHSv-IVb (MI03GL), (2) a medium dose of $4.0x10^3$ pfu/ml, (3) a high dose of $1.0x10^5$ pfu/ml, or (4) 1 ml sterile maintenance minimum essential media (MEM), which served as the negative control. The fish groups then were divided into two tank replicates (45 fish X 2), from which two fish each were selected randomly for analysis (measuring VHSv levels in their spleen tissue) at pre-determined time intervals (0, 6, 12, 24, 36 h; 2, 4, 6, 8, 15, 22, 28, 35, 42 days). Fish were euthanized immediately with an overdose of 25 mg/ml tricaine methanesulfonate (MS-222; Argent Chemical Lab), following MSU Institutional Animal Care and Use Committee approved protocols (AUF 07/07-123-00). Tanks were monitored every 8 h, and any moribund or dead fish were removed. Exterior viral particles were eliminated from the fish by submerging each 3X in double distilled H_2O. Fish were dissected under aseptic conditions with the surgical site (anus to operculum) disinfected using 100% ethanol and betadine. Spleen and head kidney were removed, placed into separate 1.5 ml tubes, flash frozen in liquid nitrogen or placed in RNAlater (Qiagen), and stored at -80^oC awaiting StaRT-PCR. Nets were disinfected using a 2% chlorhexidine solution, and sterile equipment and new gloves were used for each fish. Specimen disposal followed MSU biohazard protocols.

A second set of VHSv-IVb laboratory challenge experiments was conducted at the United States Geological Survey (USGS) Western Fisheries Research Center (WFRC) Challenge Facility in Seattle WA (under the supervision of J. Winton and F. Goetz) using six-month-old yellow perch, which were VHSv-certified-free Choptank broodstrain (Rosauer et al., 2011) from the University of Wisconsin's Great Lakes WATER Institute (Milwaukee, WI). Laboratory work was conducted under USGS-WFRC Animal Care and Use guidelines. Perch were kept in ~8 m diameter covered 278

L aquaria at 18-20^{o}C under ambient light conditions that mirrored the seasonal photoperiod, and fed 1.2 mm pellet feed (Oregon Biodiet, Longview, WA) every other day until satiation. A total of 210 perch were randomly selected and assigned into six groups: two groups of 38 fish (mean=15 g) were challenged via intra-peritoneal injection of $1.0x10^5$ pfu/ml VHSv-IVb (MI03GL), a group of 20 fish (mean=15 g) was immediately euthanized, serving as a negative control, and three groups of 38 fish each (mean=~2 g) were used for the VHSv immersion challenge, with two groups immersed for two hours in the same VHSv dosage ($1.0x10^5$ pfu/ml), and the final group challenged with a control dose of MEM-0 (as described above). Fish were randomly selected at pre-determined intervals (10 fish in day 1, eight in day 2, and five each for days 3-6) and euthanized using 240 mg/L MS-222 and 1.2g/L $NaHCO_3$ following USGS-WFRC Institutional Animal Care and Use Committee protocols (2008-17). Spleen and head kidney were removed from each fish with sterile equipment, labeled, flash frozen in liquid nitrogen, and stored at -80^{o}C. Specimens were disposed of following University of Wisconsin's Great Lakes WATER Institute and University of Toledo biohazard protocols.

4.3.6 Testing for VHSv infection using StaRT-PCR and other assays

The relative performances of StaRT-PCR, conventional SYBR® green qRT-PCR, and cell culture to detect VHSv infection were compared using a χ^2 test (Sokal and Rohlf, 1995) for 23 wild-caught fishes from the Great Lakes, including: two bluegill ((*Lepomis macrochirus*), May 2011, Budd Lake, MI), a brown bullhead ((*Ameiurus nebulosus*),

May 2012, Maumee Bay, Lake Erie), a freshwater drum (April 2012, Sandusky Bay, Lake Erie), seven largemouth bass (four from May 2011 and two from July 2011, Budd Lake, MI, and one from April 2012, Sandusky Bay, Lake Erie), one smallmouth bass (May 2006, Sodus Bay, Lake Ontario), and 11 lake herring ((*Coregonus artedi*), December 2009, Apostle Islands, Lake Superior). Other evaluations with StaRT-PCR included 20 experimentally challenged muskellunge (15 VHSv infected and five negative controls) and 20 of the challenged yellow perch (including seven infected fish and three negative controls from the immersion and injection challenge experiments). Cell culture was not performed on the perch samples due to the low tissue quantities available from the WFRC.

SYBR® green qRT-PCR was conducted in 25 µl rxns, containing 0.05 µg of each primer (primers were the same as those for StaRT-PCR; Table 4.1), 2 µl cDNA product, 10 µl SoFast SYBR® green mix, and RNase-free water. Amplifications were carried out on a Mastercycler Realplex Thermocycler (Eppendorf, Inc., Westbury, NY) using an initial denaturation of 5 min at 95^{o}C, followed by 40 cycles of 30 sec at 95^{o}C, and 1 min at 60^{o}C. Each rxn included a known cell culture positive, a negative VHSv cDNA, and a reagent negative control (nuclease-free H_2O). Positive versus negative results for VHSv were determined based on C_t, with positives resulting in a $C_t \leq 38$. Use of a C_t this high was facilitated by the high signal to background achieved with these optimized reagents. All samples were analyzed in triplicate, and the products were visualized on 1% agarose gels to confirm positive/negative results.

Cell culture was performed at MSU's Aquatic Animal Health Laboratory following standard OIE (2009) procedures. Spleen tissue from individual fish samples

was homogenized using a Biomaster Stomacher (Wolf Laboratories Ltd., Deans La York, UK) at high speed for 2 min and diluted with Earl's Salt-Based MEM (Invitrogen, Grand Island, NY) supplemented with 12 mM TRIS buffer (Sigma, St. Louis, MO), penicillin (100 IU/ml), streptomycin (100 μg/ml; Invitrogen), and amphotericin B (250 μg/ ml; Invitrogen) to produce a 1:4 dilution of the original tissues. The dilutions were centrifuged at 2000 g and the supernatants were placed into individual wells of a 24 well plate containing confluent EPC cells, MEM, and 5% fetal bovine serum. Plates were incubated at 15^{o}C for seven days and observed for the formation of cytopathic effects. A second and third passage was performed before concluding infectivity.

If results were positive for cell culture (i.e., cytopathic effects were observed), RNA was extracted from infected cells following the method used for StaRT-PCR, reverse transcribed using Affinity Script Multiple Temperature Reverse Transcriptase PCR (Stratagene, La Jolla, CA), and amplified using previously described recommended standard procedures and VHSv *N*-gene primers (Fig. 4-1) (OIE, 2009).

4.3.7 Quantitative analyses using StaRT-PCR

Mean VHSv molecules calculated (from independent triplicate StaRT-PCR experiments) were compared for the wild caught and laboratory challenged fish samples. Numbers of VHSv molecules in 18 laboratory challenged muskellunge were compared between nine fish having clinical signs of infection (e.g., gross lesions and hemorrhages) and nine without signs, which were sampled at identical time points. A non-parametric Mann-Whitney U test (using SPSS) was used to rank the relative numbers of molecules

in each fish. A χ^2 test (Microsoft Excel) was used to evaluate whether a threshold number of VHSv molecules characterized the appearance of clinical signs. A power analysis (G*Power2; Erdfelder et al., 1996) estimated the sample size needed for 95% confidence, using an effect size of 0.50 (Cohen, 1992). Differences in the number of VHSv molecules across days 1-6 of infection in the laboratory challenged juvenile yellow perch were tested using a one-way ANOVA. Additionally, numbers of VHSv molecules were compared between the immersion versus injected challenged individuals (60 of each).

4.4 RESULTS

4.4.1 Performance of the VHSv StaRT-PCR test

The StaRT-PCR test was determined to be specific for VHSv diagnosis. Specifically, StaRT-PCR VHSv results were negative (no amplification) for human viruses (Encephalomyocarditis virus and Vesicular Stomatitis virus) and for other fish viruses (i.e., Hirame Rhabdovirus, Infectious Hematopoietic Necrosis virus, Infectious Pancreatic Necrosis virus, Snakehead Rhabdovirus, and Spring Viremia of Carp virus). All tests of VHSv strains and substrains (I, Ia, II, III, IVa, IVb, IVc) were positive (Table 4.3).

Regression analyses showed that 100% of StaRT-PCR tests at concentrations ≥6 VHSv molecules in the sample had 100% true accuracy (Fig. 4-2; R^2=0.93, F=190.10, df=1, 14, p<0.0001). Experiments yielded 100% (9/9 times) amplification for dilutions of ≥6 VHSv molecules, 88% (8/9 times) for 4 and 5 molecules, 44% (4/9) for 2 molecules,

33% (3/9) for 1 molecule, and 22% (2/9) for 0.5 molecules. The number of molecules calculated from StaRT-PCR did not significantly differ from the NanoDrop measurements (Fig. 4-2; χ^2=0.47, df=15, NS).

StaRT-PCR quantitation of VHSv molecules in the dilution test series using biological samples followed a linear relationship to results from plaque assays, across a range of $0.1x10^0$-$1.0x10^3$ VHSv/10^6 *actb1* molecules (Fig. 4-3; R^2=0.98, F=1797.86, df=1, 43, p<0.001). The precision of the test was calculated as 9.57% for samples ≥11 expected VHSv/10^6 *actb1* molecules and 37.69% for ≤11 molecules.

4.4.2 StaRT-PCR controls for interference in PCR

As measured by fluorescent light units, the amount of PCR product from positive VHSv-IVb fish RNA decreased in proportion to the amount of interfering substances in the StaRT-PCR assay (Fig. 4-4). For example, as EDTA increased from 0.1-1.3 mM, the fluorescence signal for both VHSv NT and IS PCR product decreased relative to the baseline and then was eliminated at 1.4 mM or higher. However, because the number of VHSv NT molecules was measured relative to IS molecules, and because both NT and IS were affected the same way by presence of EDTA, the number of VHSv NT molecules measured was unaffected until EDTA concentration increased to 1.3 mM. Even in this condition, despite reduction of fluorescence signal to 65% of control, measured VHSv molecules/10^6 *actb1* declined only slightly (by 6%; F=4.46, df=6, 14, p=0.01) compared to the control. At concentrations ≥1.4 mM EDTA, no amplification of VHSv NT or IS occurred. Since IS was not observed, this result would be recorded as non-functioning

assay, not as a negative, which would be a false negative. When testing these samples using the same protocol but without the SMIS, VHSv amplification progressively decreased with increasing EDTA concentration, with no signal at ≥1.4 mM EDTA. This comparison indicates that lack of inclusion of the SMIS would lead to a false negative report (indicated by *).

4.4.3 Optimization of reverse transcription conditions

In tests for possible RNA interference, VHSv and two reference genes, *actb1* and *ef1a*, were measured along with the RTSM across RNA concentrations from 1 (the baseline control), 5, 10, 20, and 30 μg RNA/90 μl rxn (Fig. 4-5). The reverse transcription efficiency as measured by the RTSM (Fig. 4-5a), was significantly reduced with increasing RNA input (F=9.05, df=4, 10, p=0.002). Differences were measured using pairwise tests between yields at 1 and 10, 20, and 30, and 10 versus 30 μg (t=3.10-6.20, df=4, p=0.003-0.04). Overall, reverse transcription efficiency decreased 23-26% from the baseline at 20 and 30 μg RNA/90 μl rxn, respectively. Measured numbers of both reference genes increased in parallel with RNA concentration (Fig. 4-5b; *actb1*: F=96.45, df=4, 10, p<0.001; *ef1a*: F=27.36, df=4, 10, p<0.001), yet remained constant when *ef1a* was normalized to *actb1* (F=1.01, df=4, 10, NS). Measured numbers of VHS molecules, in contrast to what we predicted, increased significantly increasing RNA input from 1 to 10 μg (Fig. 4-5c; t=2.82-3.46, df=4, p=0.03-0.05), but showed no further increase with additional RNA/reverse transcription (t=0.32-1.68, df=4, NS). The number of VHSv molecules per 106 *actb1* molecules significantly decreased with addition of

RNA/reverse transcription (F=282.64, df=4, 10, p<0.001).

4.4.4 VHSv Detection and quantitation in wild caught and laboratory challenged fishes

StaRT-PCR had greater accuracy than conventional SYBR® green qRT-PCR, with the latter having 40% false negative error in experiments with wild caught fishes (Fig. 4-6a; 10 vs. 6 positives; χ^2=1.53, df=1, NS), 47% error for immersion challenge muskellunge (Fig. 4-6b; 15 vs. 8; χ^2=5.01, df=1, p=0.03), and 14% error for challenged yellow perch (Fig. 4-6c; 14 vs. 12; χ^2=0.44, df=1, NS). StaRT-PCR also detected significantly more VHSv positives than did cell culture, with the latter having 70% false negative error for wild caught fishes (Fig. 4-6a; 10 vs. 3; χ^2=10.39, df=1, p=0.001), and 47% error in laboratory challenged muskellunge (Fig. 4-6b; 15 vs. 8; χ^2=5.01, df=1, p=0.03).

SYBR® green qRT-PCR correctly detected slightly more positives than did cell culture in the wild caught fishes (Fig. 4-6a; 6 vs. 3, 50% error difference; χ^2=1.24, df=1, NS); both diagnosed equivalent numbers of positives in the experimentally challenged muskellunge (Fig. 4-6b; 8 vs. 8; χ^2=0.00, df=1, NS). All positives detected by cell culture and SYBR® green qRT-PCR were also positive with StaRT-PCR. The false negative range for SYBR® green qRT-PCR was $1.0x10^0$-$6.5x10^1$ VHSv molecules (as quantified by StaRT-PCR) and $1.0x10^0$-$1.0x10^3$ molecules for cell culture. For all assays, negative controls yielded negative results; i.e., they had no false positives and no contamination.

Numbers of VHSv molecules measured by StaRT-PCR varied widely among wild caught specimens ($1.0x10^0$-$1.2x10^5$ VHSv molecules/10^6 *actb1* molecules) and laboratory

challenged individuals ($1.0x10^0$-$8.4x10^5/10^6$ *actb1* molecules). Fig. 4-7 shows a difference in the mean numbers of VHSv molecules between muskellunge exhibiting clinical signs of infection ($1.9x10^4 \pm 1.2x10^4$) versus those without ($1.1x10^3 \pm 4.5x10^2$; Z=-2.10, U(df)=1, p=0.04, days 6-28). At day 35, no VHSv was detected in remaining fish from either group. Numbers of VHSv molecules appeared to differ at days 6, 9, 15, and 28, with those showing clinical signs of infection having higher values (Fig. 4-7). One hundred VHSv molecules were estimated to distinguish a threshold at which individuals showed clinical signs of infection (χ^2=0.09, df=1, NS). Power calculations determined that 52 fish samples (26 with and 26 without clinical signs) would be needed to verify this level (95% confidence interval) in further testing.

Quantities of VHSv measured in the yellow perch laboratory challenge experiments ranged from $4.0x10^0$-$1.3x10^5$ VHSv molecules/10^6 *actb1* molecules in the immersion challenged fish and $1.0x10^0$-$1.8x10^5$ molecules in the injection challenged individuals. Results of the yellow perch laboratory challenge experiments showed that VHSv infection first was detected at day three in immersion challenged fish versus day two in injection challenged individuals (Fig. 4-8; t=2.15, df=19, p=0.04). Overall numbers of VHSv molecules appeared relatively consistent across the remaining course of infection through day six (immersion: F=0.48, df=3, 29, NS; injection: F=2.62, df=4, 42, p=0.05), and did not differ between the two experiments (t=0.03-0.96, df=13-18, NS).

4.5 DISCUSSION

4.5.1 Specificity and Accuracy of the StaRT-PCR VHSv test (Hypotheses 1 and 2)

The fisheries, aquaculture, and baitfish industries rely on accurate certification of their products as VHSv-free (Aquatic Invasive Species Action Plan, 2011), for which results from cell culture may take up to 28 days (Garver et al., 2011) and frequently yield false negative conclusions (47-70% reported in this study). Thus, the development of a rapid and accurate diagnostic assay is key to preventing VHSv spread to new areas, species, and populations (Cutrín et al., 2009). Similar to other PCR-based assays developed to detect VHSv, the StaRT-PCR test is VHSv specific, supporting hypothesis H_A1. StaRT-PCR detected all VHSv strains and substrains, and did not amplify other viruses.

Antibody assays (Millard and Faisal, 2012) and other PCR tests (Chico et al., 2006; López-Vázquez et al., 2006; Liu et al., 2008; Matejusova et al., 2008; Cutrín et al., 2009; Hope et al., 2010; Garver et al., 2011; Jonstrup et al., 2012; Phelps et al., 2012), lacked internal controls. Those other PCR tests often resulted in false negatives, leading to inaccurate conclusions, as demonstrated by the present investigation's SYBR® green qRT-PCR results (14-47% false negative error).

4.5.2 Quantitation of VHSv by StaRT-PCR (Hypothesis 2)

Results support hypothesis H_A2 that StaRT-PCR accurately quantifies VHSv and

is effective across all levels of the virus tested, from low to high. Other assay methods were unable to discern lower levels of virus, resulting in false negatives. Notably, cell culture was unable to discern levels from $1.0x10^{0}$-$1.0x10^{3}$ VHSv/10^{6} *actb1* molecules and SYBR® green qRT-PCR failed to detect from $1.0x10^{1}$-$6.5x10^{1}$ molecules. Fish harboring low levels of the virus thus would pass inspection for VHSv-free certification, using those assays.

PCR-based methods have been shown to be more sensitive than traditional cell culture (e.g., Chico et al., 2006; López-Vázquez et al., 2006; Hope et al., 2010; Garver et al., 2011; Jonstrup et al., 2012) since they amplify replicating (i.e., infectious) and non-replicating (i.e., non-active) transcripts alike. Cell culture solely detects actively replicating virus that is capable of infecting cells. Thus, it was anticipated that StaRT-PCR, like other RT-PCR approaches, would yield more positives than cell culture. However, detection of both replicating and non-replicating VHSv RNA could be advantageous since its presence may indicate that there are active infections in the area where the samples were collected. This information may aid management efforts and aquaculture facilities to identify latent infections that could be negative by cell culture, but potentially infective.

StaRT-PCR detected more positives than SYBR® green qRT-PCR, which was not surprising as it was more thoroughly optimized for PCR efficiency and controlled for false negatives. Compared to other PCR tests, StaRT-PCR detected a lower threshold of molecules. For example, Garver et al. (2011) stated that their two-step qRT-PCR assay diagnosed 100 VHSv N-gene copies with 100% accuracy. In contrast, StaRT-PCR showed 100% accuracy for six molecules and detected samples of a single molecule at a

frequency limited only by the laws of chance governing stochastic sampling variation.

4.5.3 Performance in the presence of PCR inhibitors (Hypothesis 3)

StaRT-PCR effectively controlled for inhibition at the PCR level, supporting H_A3. In contrast, Hope et al.'s (2010) one-step qRT-PCR test showed that increasing the VHSv concentration in reverse transcription rxns from 50 ng to 1 µg RNA/25 µL resulted in only a 10-fold increase in PCR product, two-fold less than what they expected, which could lead to false negatives. They observed even more dramatic reduction in reverse transcription efficiency with higher RNA input. Based on our results, we conclude that 1 µg/90 µl reverse transcription rxn will provide a test with the best balance between optimal sensitivity and adverse effects of RNA input on reverse transcription efficiency. Further increase in RNA input to 10 µg/90 µl will yield further significant increase in VHSv amplification, but the gain will be marginal relative to the RNA consumed due to decreasing reverse transcription efficiency.

Degraded or environmentally challenged samples can lead to loss of signal and interfere with detection of the virus (see McCord et al., 2011). Such samples often contain substances that are co-extracted with the RNA or carried over into the cDNA via the reverse transcription rxn, which may inhibit or lower PCR efficiency by binding to the polymerase and/or blocking reagents necessary for amplification (Opel et al., 2010). StaRT-PCR is the sole VHSv diagnostic to test, detect, and control for PCR inhibition, preventing false negative results when samples contained known chemical inhibitors.

The amount of VHSv PCR product decreased as the reverse transcription

efficiency decreased, in contrast to the increase in product yield from the two reference genes, indicating that detection of this virus may be affected by RNA interference differently than the endogenous controls. The difference between them might be due to contaminating VHSv protein features that are tightly bound and not removed during the RNA extraction procedure. Antigenic sites or secondary structures (α-helicase, β-strands and loops), which are common in Novirhabdoviruses (Walker and Kongsuwan, 1999), may contribute to this difference. Other unknown factors may be responsible, which merit further investigation. To avoid these issues, a standard RNA input/reverse transcription condition should be employed when assaying for VHSv, for which 1 μg RNA/90 μl rxn is recommended.

4.5.4 StaRT-PCR performance versus other assays (Hypotheses 4 and 5)

StaRT-PCR results diagnosed and quantified the amount of VHSv in fish samples from the field and laboratory experiments, supporting H_A4. Like most other PCR-based VHSv detection methods (Chico et al., 2006; López-Vázquez et al., 2006; Hope et al., 2010), StaRT-PCR is more sensitive and accurate than the traditional "gold-standard" cell culture approach in detecting qualitative positives. In contrast to the present findings, Jonstrup et al.'s (2012) found that cell culture significantly exceeded the detection ability of their one-step qRT-PCR test at low titer ($1.9x10^2$-$1.9x10^3$ $TCID_{50}$/ml). Thus, their test, unlike StaRT-PCR, did not outperform cell culture.

Besides the present investigation, only two other studies compared detection levels of different PCR assays. López-Vázquez et al. (2006) found that their two-step

nested qRT-PCR assay had 15-80% false negatives, their qualitative one-step qRT-PCR assay had 60-90%, and cell culture had 95%. Chico et al. (2006) reported that their two-step TAQman® qRT-PCR test had an error rate of 25% (detected 9 of 12 positives) versus 92% error in two-step nested qRT-PCR assays (found 1 of 12 positives), and 67% error with cell culture (identified 4 of 12 positives). In contrast, StaRT-PCR in the present study yielded 0 false negatives, versus 14-47% false negatives with two-step SYBR® green qRT-PCR, and 47-70% with cell culture. The error rates found with SYBR® green qRT-PCR and cell culture in the present study thus were similar to those determined by Chico et al. (2006).

Most other PCR methods conducted a dilution series with a regression analysis to evaluate sensitivity (e.g., Chico et al., 2006; López-Vázquez et al., 2006; Matejusova et al., 2008; Cutrín et al., 2009; Hope et al., 2010; Garver et al., 2011; Jonstrup et al., 2012; Phelps et al., 2012). This also was done in the present study, along with an extreme limiting dilution assay that compared the observed frequency of positive StaRT-PCR assays against those predicted by Poisson distribution. Thus, the current investigation evaluated the absolute accuracy of test results, incorporating stochastic sampling variation. The close agreement between the values measured with spectrophotometric quantitation and those expected from the Poisson distribution supports the reliability of StaRT-PCR quantitation capability (Vogelstein and Kinzler, 1999). Based on this analysis, the StaRT-PCR assay measures to a single VHSv molecule with known true accuracy, unlike other PCR methods (Liu et al., 2008; Hope et al., 2010; Garver et al., 2011), whose minimum detection thresholds were much higher (requiring ≥100 copies of VHSv). Thus, StaRT-PCR has greater sensitivity and accuracy than cell culture or other

qRT-PCR tests, supporting H_A5.

4.5.5 Biological levels of VHSv detected with StaRT-PCR (Hypotheses 6-8)

The application of employing StaRT-PCR to quantify VHSv infection was demonstrated in a variety of biological experiments. Results show that levels of VHSv molecules in wild caught fishes vary widely ($1.0x10^0$-$1.2x10^5$ VHSv/10^6 *actb1* molecules), resembling values measured in laboratory challenged individuals ($1.0x10^0$-$8.4x10^5$). Additionally, mean VHSv levels in laboratory challenged individuals showing signs of VHSv infection were higher than without, supporting H_A6. This suggests that individuals in a population respond variably to VHSv infection, underscoring the importance of an improved diagnostic test for managing this disease. A threshold level of 100 VHSv molecules was identified as a potential biomarker for clinical signs of infection. This observation supports the conclusion that the greater sensitivity of StaRT-PCR compared to cell culture or SYBR® green qRT-PCR has biological significance. Further experiments are warranted to validate this finding.

To the authors' knowledge, only a single other investigation quantified VHSv samples from infected fish samples. Similar to StaRT-PCR results, Chico et al. (2006) determined that VHS viral load varied widely among nine rainbow trout individuals in an immersion challenge (measured as VHSv relative to RNA). Chico et al. (2006) was one of four other studies that have used a reference gene (*18srRNA*) for quantitation. Matejusova et al. (2008), Garver et al. (2011), and Jonstrup et al. (2012) employed a different reference gene (*ef1a*) to confirm RNA quality for cDNA synthesis and to serve

as a positive control, but did not quantify amounts of VHSv. The present study evaluated both of those reference genes (*18srRNA* and *ef1a*), as well as *actb1*, facilitating accurate quantitation of VHSv.

Over recent years, increasing numbers of individuals and species of Great Lakes fishes have tested positive for VHSv-IVb, yet many of those have appeared healthy (Kim and Faisal, 2010). For example, a freshwater drum collected in the Lake Erie Harbor of Sandusky Bay on April 12, 2012 and a largemouth bass on May 10, 2012 that tested positive for VHSv with StaRT-PCR ($3.4x10^2$ and $5.9x10^2$ VHSv/10^6 *actb1* molecules, respectively), did not have lesions, but were swimming erratically. Lack of external VHS hemorrhages renders it difficult to determine whether fish have VHS and potentially could transmit the virus. As described above, our results in experimentally challenged fish suggest that a measured value of greater than 100 using StaRT-PCR may be associated with clinical signs of infection.

No other study to date has monitored levels of VHSv molecules across the early course of infection. The relative number of VHSv molecules was tracked in experimentally challenged yellow perch from days one to six. Onset of VHSv infection occurred a day later in immersion challenged than in injection challenged fish, with their viral concentrations remaining relatively consistent through day six, supporting H_O7. Quantities of virus did not differ between the immersion versus injection challenged fish after onset of infection (H_O8). Both experiments showed a linear increase in the number of VHSv molecules from the onset of infection, followed by a plateau phase, and a decrease at day six. These data may indicate that VHSv levels are highest at the beginning of the infectious stage, level off as the host develops its immune response, and

then start to decline. However, these results are based solely on early infection, so future studies that include middle and later stages are needed to better understand VHSv infection patterns, the disease course, and host response.

4.5.6 Summary and conclusions

The cell culture diagnostic approved for VHSv-free certification (Aquatic Invasive Species Action Plan, 2011) is lengthy, labor intensive, and lacks sensitivity compared to PCR-based assays. Reliance on cell culture alone could result in spread of VHSv throughout aquaculture systems, baitfish transport, and/or watersheds, leading to significant losses. The other PCR-based techniques to detect VHSv (Chico et al., 2006; López-Vázquez et al., 2006; Liu et al., 2008; Matejusova et al., 2008; Cutrín et al., 2009; Hope et al., 2010; Garver et al., 2011; Jonstrup et al., 2012; Phelps et al., 2012) lacked the intrinsic quality control necessary for accurate and reliable detection. Moreover, those tests were unable to measure their accuracy and reverse transcription efficiency. The new StaRT-PCR assay incorporates IS that allow for sensitive and accurate qualitative and quantitative measurements of VHSv. These standards guard against false negative results and correct for interfering substances. In addition, the present investigation uses StaRT-PCR to demonstrate reverse transcription efficiency and determine a biological threshold level for clinical signs of infection. Implementation of the StaRT-PCR test will aid aquaculture, baitfish, and fishery industries via faster, more sensitive, and accurate disease detection. Application may lead to improved management and commerce, and lead to cost-savings for stakeholders. This StaRT-PCR diagnostic will enhance natural

resource conservation efforts to detect the virus and track its spread.

4.6 ACKNOWLEDGEMENTS

This study was funded by grants to C. Stepien, J. Willey, and D. Leaman from USDA-NIFA (CSREES) #2008-38927-19156, #2009-38927-20043, #2010-38927-21048, USDA-ARS CRIS project #3655-31320-002-00D, under the specific cooperative agreement #58-3655-9-748 A01 (for which B. Shepherd is another PI), and NOAA Ohio Sea Grant #R/LR-015 (to C. Stepien and J. Willey). NSF GK-12 #DGE-0742395 to C. Stepien (PI) paid the stipend of L. Pierce. The Great Lakes Fishery Trust, Lansing, MI, and the National Institutes of Health provided funds for M. Faisal and R. Kim. Special thanks are extended for T. Blomquist, A. Pore, J. Yeo, present and past members of the Great Lakes Genetics/Genomics Laboratory, including H. Dean, A. Haponski, S. Karsiotis, D. Murphy, M. Neilson, C. Prichard, O.J. Sepulveda-Villet, T. Sullivan, and S. Woolwine, and administrative staff P. Uzmann and M. Gray. M. Salit of the National Institutes of Standards and Technology (Gaithersburg, MD) provided the ERCC standards to access reverse transcription efficiency. This study was aided by samples from G. Kurath and J. Winton of USGS (Seattle, WA), P. Bowser of Cornell University (Ithaca, NY), F. Goetz of the University of Wisconsin (Milwaukee, WI), K. Garver of Fisheries and Oceans (Ottawa, Canada), I. Bandín of the Universidad de Santiago de Compostela (La Coruña, Spain), and T. Gadd of the Finnish Food Safety Authority Evira (Helsinki, Finland). The views contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Government. Mention of trade name, proprietary product, or

specific equipment does not constitute a guarantee or warranty by the USDA and does not imply its approval to the exclusion of other products that may be suitable. This manuscript is submitted for publication with the understanding that the U.S. Government is authorized to reproduce and distribute reprints for governmental purposes. This is publication #2013-07 from the Lake Erie Research Center.

Table 4.1

Sequences and PCR parameters of primers and internal standards (IS) used for StaRT-PCR. F=forward, R=reverse, CT=competitive template.

Primer and IS	Location	Sequence (5'-3')	T_m (°C)	GC %	Product length (bp)
N-gene:					
VHSv F5	362-383	GTC CGT GCT TCT CTC CTA TGT	66.9	52.4	273
VHSv R5	614-635	TCC CCG AGT TTC TTA GTG ATG	67.6	47.6	273
Reference Gene Primers:					
actb1 F1	620-640	TTG GCT GGC CGT GAC CTC AC	77.6	65.0	190
actb1 R1	790-810	GCA GCT CGT AGC TCT TCT CC	67.9	60.0	190
ef1a F1	239-259	ACC ACC GGC CAT CTG ATC TA	71.2	55.0	220
ef1a R1	440-459	TGT GTC CAG GGG CAT CAA T	70.9	52.6	220
18srRNA F1	4513-4533	AGT ACA CAC GGC CGG TAC AG	69.5	60.0	269
18srRNA R1	4762-4786	GGG CAG ACA TTC GAA TGA GA	69.6	50.0	269
Reverse Transcription Efficiency Primers:					
ERCC 113 F	116-136	TTG GAT CAG TGG GAA GTG CT	69.1	50.0	130
ERCC 113 R	226-246	GGG GCT CGA AAG GTA CTA GG	68.5	60.0	130
ERCC 171 F	153-173	AAG CTG ACG GTG ACA AGG TT	68.2	50.0	118
ERCC 171 R	251-271	TCG CAG TTT TCC TCA AAT CC	68.6	45.0	118
Internal Standards:					
VHSv CT5	543-585	TCC CCG AGT TTC TTA GTG ATG CAA GGT CCC CTT GAC GAT TTC	89.9	50.0	223
actb1 CT1	747-787	GCA GCT CGT AGC TCT TCT CCG CCC ATC TCC TGC TCG AAG T	91.9	60.0	167
ef1a CT1	389-427	TGT GTC CAG GGG CAT CAA TTC CAG AGA GCG ATG TCG AT	91.5	52.6	188
18srRNA CT1	4671-4711	GGG CAG ACA TTC GAA TGA GAA TGA GAG AGG CAC CCC GCA TGG GTT T	94.3	57.5	198
ERCC 113 CT	116-218	TTG GAT CAG TGG GAA GTG CTC ACG CGC GGA GCC CAC TGG GCG AAC AGC AAC GTT ATA ACG GCC ACT CAG TGG TTC GTC ACG CCC TAG TAC CTT TCG AGC CCC	87.0	59.8	102
ERCC 171 CT	153-250	AAG CTG ACG GTG ACA AGG TTT CCC CCT AAT CGA GAC GCT GCA ATA ACA CAG GGG CAT ACA GTA ACC AGG CAA GAG TTG GAT TTG AGG AAA ACT GCG A	82.5	49.5	97

Table 4.2

StaRT-PCR SMIS (Standardized Mixture of Internal Standards) dilution mixtures of the four IS (internal standards): *18srRNA/actb1/ef1a/VHSv*. Values are reported as 10^X M.

A	B	C	D	E	F	G	H
-10/-11/-10/-10	-10/-11/-10/-11	-10/-11/-10/-12	-10/-11/-10/-13	-10/-11/-10/-14	-10/-11/-10/-15	-10/-11/-10/-16	-10/-11/-10/-17
-11/-12/-11/-11	-11/-12/-11/-12	-11/-12/-11/-13	-11/-12/-11/-14	-11/-12/-11/-15	-11/-12/-11/-16	-11/-12/-11/-17	
-12/-13/-12/-12	-12/-13/-12/-13	-12/-13/-12/-14	-12/-13/-12/-15	-12/-13/-12/-16	-12/-13/-12/-17		
-13/-14/-13/-13	-13/-14/-13/-14	-13/-14/-13/-15	-13/-14/-13/-16	-13/-14/-13/-17			
-14/-15/-14/-14	-14/-15/-14/-15	-14/-15/-14/-16	-14/-15/-14/-17				
-15/-16/-15/-15	-15/-16/-15/-16	-15/-16/-15/-17					
-16/-17/-16/-16	-16/-17/-16/-17						
-16/-17/-16/-17							

Table 4.3

Viral isolates screened using StaRT-PCR. -=negative result (no amplification), +=positive result.

Type	Isolate	Result
Human:		
Encephalomyocarditis virus		-
Vesicular Stomatitis virus		-
Fish:		
Hirame rhabdovirus[a]		-
Infectious Hematopoietic Necrosis virus (strain 220-90)[a]		-
Infectious Pancreatic Necrosis virus		-
Snakehead rhabdovirus[a]		-
Spring Viremia of Carp virus[a]		-
VHSv:		
I	DK-F1[b]	+
Ia	FR0771[b]	+
Ia	JP96KRRV9601[b]	+
II	FI-ka663-06[c]	+
III	SM2897[d]	+
III	SC2645[d]	+
III	GH35[d]	+
III	GH 44[d]	+
IVa	Bogachiel[b]	+
IVa	Cod'91[b]	+
IVa	Elliott Bay[b]	+
IVa	JP96Obama[b]	+
IVa	Makah[b]	+
IVa	Orcas[b]	+
IVb	MI03GL[a,b]	+
IVb	vcG002[a]	+
IVb	vcG003[a]	+
IVb	vcG004[a]	+
IVb	vcG005[a]	+
IVb	vcG006[a]	+
IVb	vcG007[a]	+
IVb	vcG008[a]	+
IVb	vcG009[a]	+
IVb	vcG010[a]	+
IVc	CA-NB00-02[e]	+

Isolates obtained from:
[a]Western Fisheries Research Center, USGS, Seattle, WA, USA
[b]Cornell University College of Veterinary Medicine, Ithaca, NY, USA
[c]Finnish Food Safety Authority Evira, Finland
[d]Universidad de Santiago de Compostela, Spain
[e]Fisheries and Oceans Canada, Pacific Biological Station, BC, Canada

Fig. 4-1 Nucleotide map of the VHSv *N*-gene, showing locations of our StaRT-PCR primers in relation to those from other studies (Chico et al. 2006; Cutrín et al. 2009; Garver et al. 2011; Hope et al. 2010; Jonstrup et al. 2012; López-Vázquez et al. 2006; Matejusova et al. 2008; OIE 2009).

Fig. 4-2 Relationship between the numbers of VHSv positives discerned by StaRT-PCR (based on the % from nine separate runs of a known positive fish, for 16 dilutions of VHSv molecules) versus the number expected, as calculated via a Poisson distribution (R^2=0.93, F=190.10, df=1, 14, p<0.0001). Dilutions were: 0.5, 1, 2, 3, 4, 5, 6, 8, 10, 25, 50, $1.0x10^2$, $2.0x10^2$, $6.0x10^2$, $6.0x10^3$, and $6.0x10^6$ VHSv molecules. Results showed that all eight dilutions ≥6, yielded 100% positives using StaRT-PCR (thus their ratio was 1.0 between observed versus expected). Values under the points are the number of VHSv molecules measured by NanoDrop. The number of molecules detected by StaRT-PCR did not statistically differ from the NanoDrop values (χ^2=0.47, df=15, NS).

Fig. 4-3 The numbers of VHSv molecules determined from StaRT-PCR versus those from plaque assay (as $\log_{10}$) across a known dilution series of 10^0-10^5 pfu VHSv-IVb (strain MI03GL)/10^6 EPC cells. Closed circle=mean of three replicate runs and three replicate measures each (nine total), SE=solid line, range=dotted line. The relationship was linear: R^2=0.98, F=1797.86, df=1, 43, p<0.001.

Fig. 4-4 The relationship between the numbers of VHSv molecules determined from StaRT-PCR versus the EDTA concentration (a test for possible inhibition). Closed

circle=mean number of VHSv molecules per 10^6 *actb1* molecules from three replicates ± SE (these changed by only 6% from 0-1.3 mM EDTA). Open circle=mean fluorescent light units from three replicates ± SE (fluorescent light units decreased while the number of VHSv molecules remained relatively constant; F=4.46, df=6, 14, p=0.01). *=no StaRT-PCR amplification occurred in either the NT or the IS (thus, the rxn at this point was inhibited).

Fig. 4-5 Mean reverse transcription efficiency and performance of StaRT-PCR under possible RNA inhibition conditions (± SE from three replicates). (a) The relationship between the mean numbers of ERCC 171 molecules/10^6 ERCC 113 molecules from StaRT-PCR versus the concentration of VHSv positive RNA used in the rxn (μg RNA/90 μl). Reverse transcription efficiency significantly changed from 1 (the baseline control), 5, 10, 20, and 30 μg RNA/90 μl rxn (F=9.05, df=4, 10, p=0.002), with significant pairwise tests at 1 and 10, 20, and 30, and 10 versus 30 μg (t=3.10-6.20, df=4, p=0.003-0.04). (b) The relationship between the number of molecules from StaRT-PCR using the reference genes *actb1* and *ef1a* versus the concentration of RNA used in the rxn (μg RNA/90 μL). The numbers of *actb1* and *ef1a* molecules both increased with RNA concentration (*actb1*: F=96.45, df=4, 10, p<0.001; *ef1a*: F=27.36, df=4, 10, p<0.001), but remained relatively constant (dotted line) when *ef1a* was normalized to *actb1* (reported as $\log_{10}$; F=1.01, df=4, NS). (c) The relationship between the mean number of molecules from StaRT-PCR, using the target gene VHSv and the reference gene *actb1*, versus the concentration of RNA used in the rxn (μg RNA/90 μL). The number of VHSv molecules per 10^6 *actb1* molecules significantly decreased (dotted line) as more RNA/reverse

transcription was added (reported as $\log_{10}$; F=282.64, df=4, 10, p<0.001), suggesting that RNA concentrations >1 μg RNA/90 μl may result in inaccurate VHSv quantitation.

Fig. 4-6 Comparisons among assays test results for VHSv infection from StaRT-PCR, SYBR® green qRT-PCR, and cell culture for (a) Wild caught fishes, (b) Muskellunge challenge experiments, and (c) Yellow perch challenge experiments (cell culture not available). Compared with StaRT-PCR, for (a) SYBR® green qRT-PCR had 40% false negative error (χ^2=1.53, df=1, NS) and cell culture had 70% error (χ^2=10.39, df=1, p=0.001). For (b), SYBR® green qRT-PCR and cell culture both had 47% false negative error (χ^2=5.01, df=1, p=0.03). For (c), SYBR® green qRT-PCR had 14% false negative error (χ^2=0.44, df=1, NS). True positives=white, true negatives=gray, false negatives=gray with hash marks.

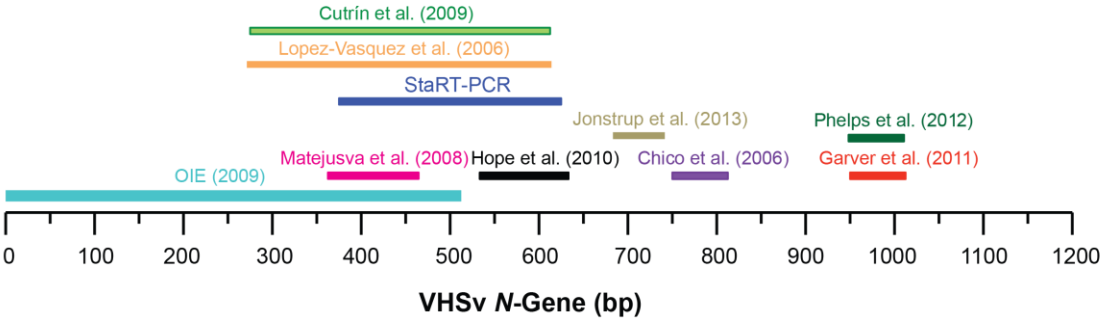
Fig. 4-7 Comparative numbers of VHSv molecules per 10^6 *actb1* molecules ($\log_{10}$ mean ± SE of three replicate runs) in laboratory challenged muskellunge with (closed circle) and without (open circle) clinical signs of VHSv infection, measured using StaRT-PCR. Values for the two groups statistically differed in a Mann Whitney ranking test (Z=-2.10, U(df)=1, p=0.04 for days 6-28).

Fig. 4-8 Numbers of VHSv molecules per 10^6 *actb1* molecules measured by StaRT-PCR (mean ± SE from three replicates; reported as $\log_{10}$) across the early stages of infection in yellow perch for (a) Injection challenge and (b) Immersion challenge experiments. Closed circles=overall mean. Open squares=individual measures. The experiments

differed in the onset of infection, with (a) being three days and (b) being two (t=2.15, df=19, p=0.04). Overall numbers of VHSv molecules remained consistent across the infection course ((a): days 3-6, F=0.48, df=3, 29, NS, (b): days 2-6, F=2.62, df=4, 42, p=0.05). There was no difference in the number of VHSv molecules between (a) and (b) (t=0.03-0.96, df=13-18, NS).

Fig. 4-1

(a)


Cutrín et al. (2009)
Lopez-Vasquez et al. (2006)
StaRT-PCR
Jonstrup et al. (2013)
Phelps et al. (2012)
Matejusva et al. (2008)
Hope et al. (2010)
Chico et al. (2006)
Garver et al. (2011)
OIE (2009)
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
VHSv N-Gene (bp)

(b)

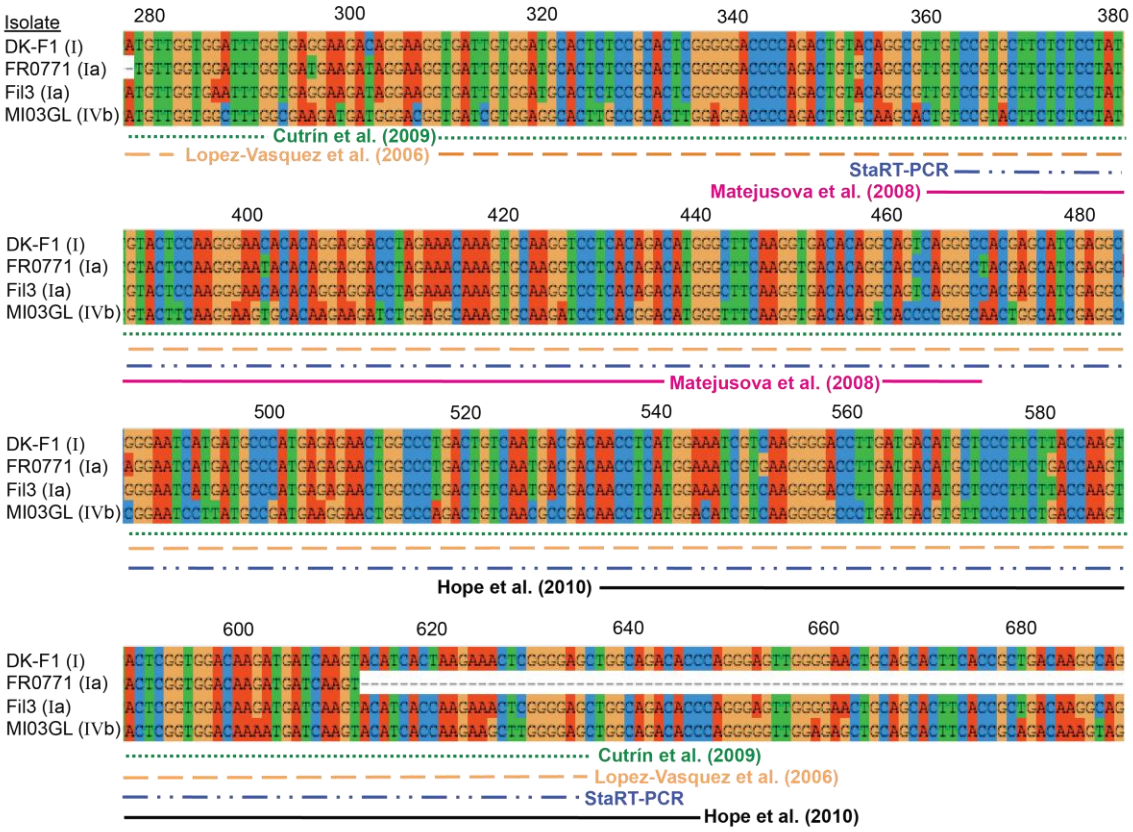

Isolate
DK-F1 (I)
FR0771 (Ia)
Fil3 (Ia)
MI03GL (IVb)
Cutrín et al. (2009)
Lopez-Vasquez et al. (2006)
StaRT-PCR
Matejusova et al. (2008)
Hope et al. (2010)

Fig. 4-2

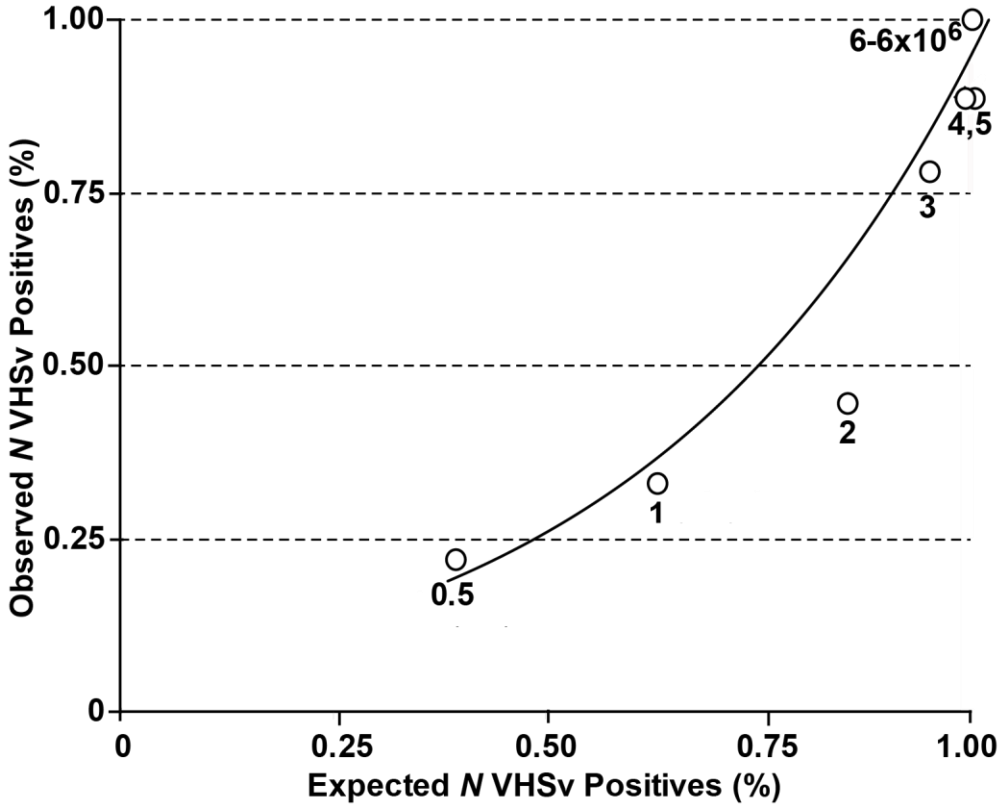

1.00
0.75
0.50
0.25
0
0
0.25
0.50
0.75
1.00
Observed N VHSv Positives (%)
Expected N VHSv Positives (%)
6-6x10^6
4,5
3
2
1
0.5

Fig. 4-3

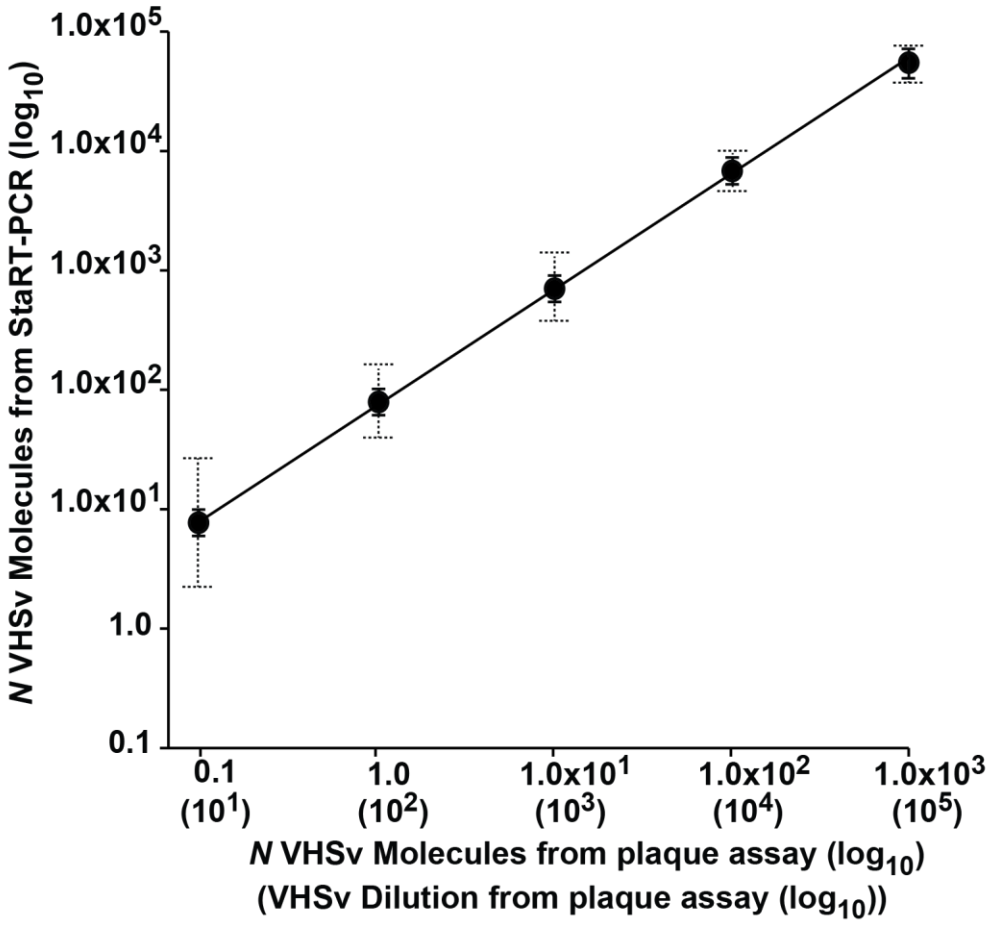

1.0x10^5
1.0x10^4
1.0x10^3
1.0x10^2
1.0x10^1
1.0
0.1
N VHSv Molecules from StaRT-PCR (log10)
0.1 (10^1)
1.0 (10^2)
1.0x10^1 (10^3)
1.0x10^2 (10^4)
1.0x10^3 (10^5)
N VHSv Molecules from plaque assay (log10)
(VHSv Dilution from plaque assay (log10))

Fig. 4-4

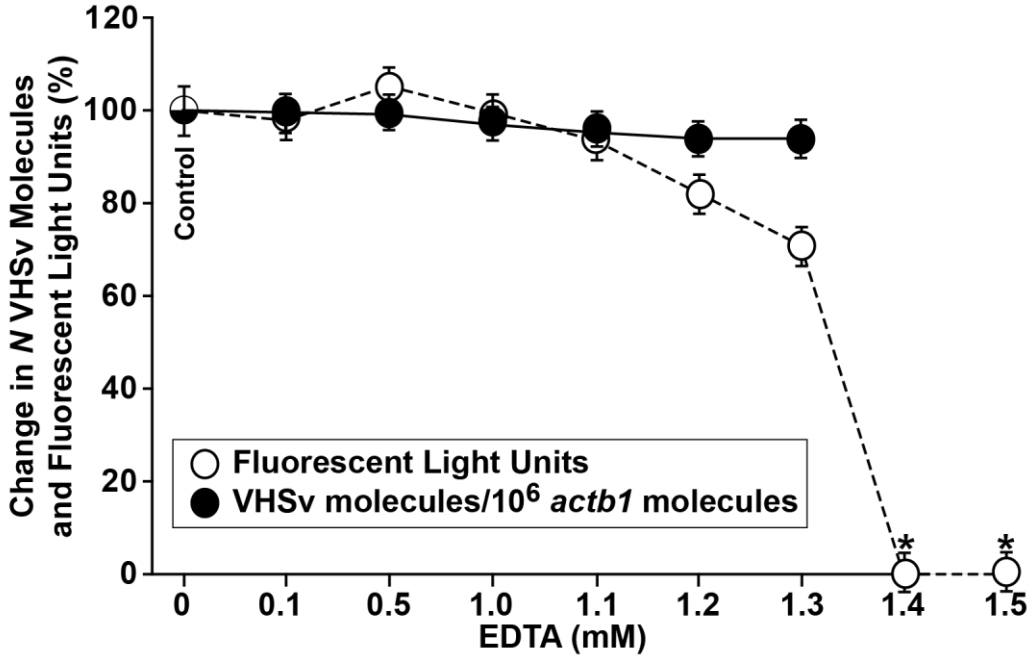

Change in N VHSv Molecules and Fluorescent Light Units (%)
120
100
80
60
40
20
0
Control I
Fluorescent Light Units
VHSv molecules/10^6 actb1 molecules
*
*
0
0.1
0.5
1.0
1.1
1.2
1.3
1.4
1.5
EDTA (mM)

Fig. 4-5

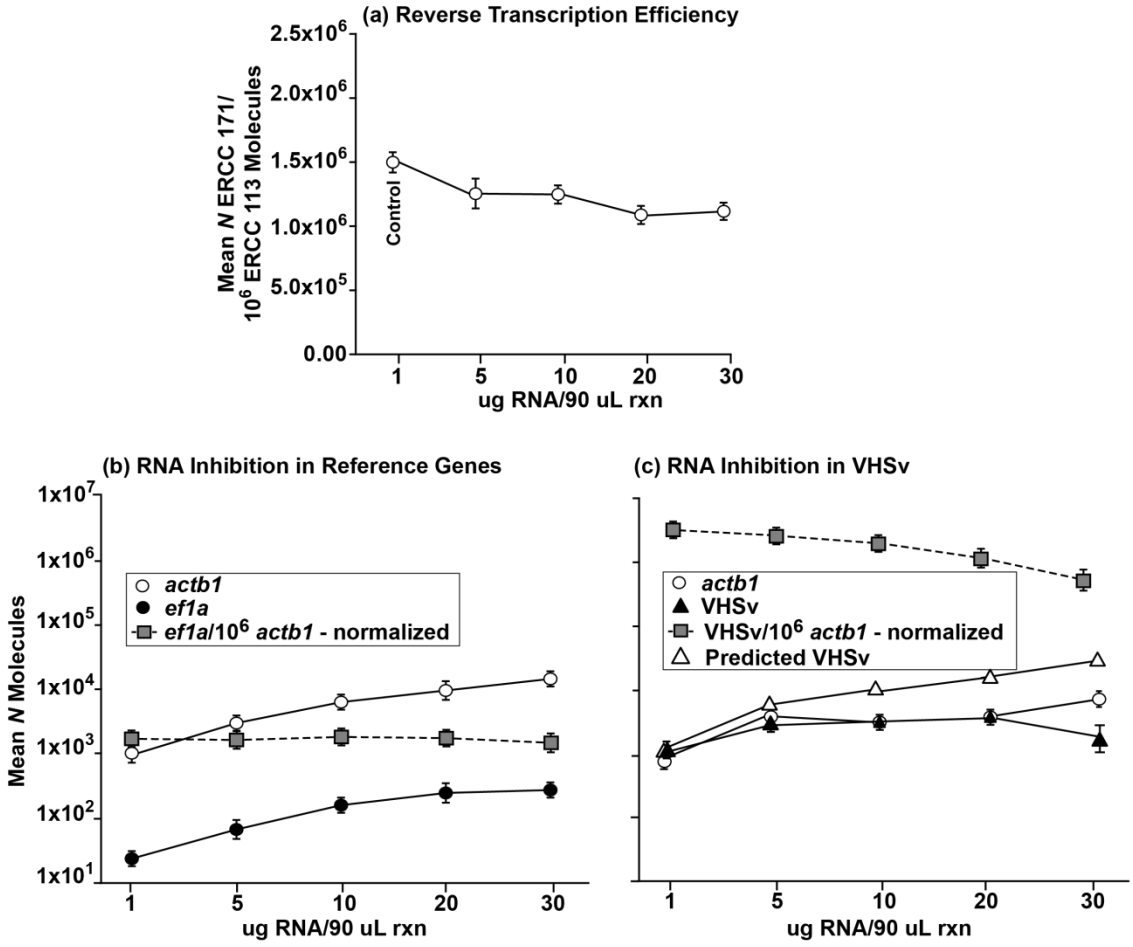

(a) Reverse Transcription Efficiency
Mean N ERCC 171/ 10^6 ERCC 113 Molecules
2.5x10^6
2.0x10^6
1.5x10^6
1.0x10^6
5.0x10^5
0.00
Control RT
1
5
10
20
30
ug RNA/90 uL rxn
(b) RNA Inhibition in Reference Genes
Mean N Molecules
1x10^7
1x10^6
1x10^5
1x10^4
1x10^3
1x10^2
1x10^1
actb1
ef1a
ef1a/10^6 actb1 - normalized
1
5
10
20
30
ug RNA/90 uL rxn
(c) RNA Inhibition in VHSv
actb1
VHSv
VHSv/10^6 actb1 - normalized
Predicted VHSv
1
5
10
20
30
ug RNA/90 uL rxn

Fig. 4-6

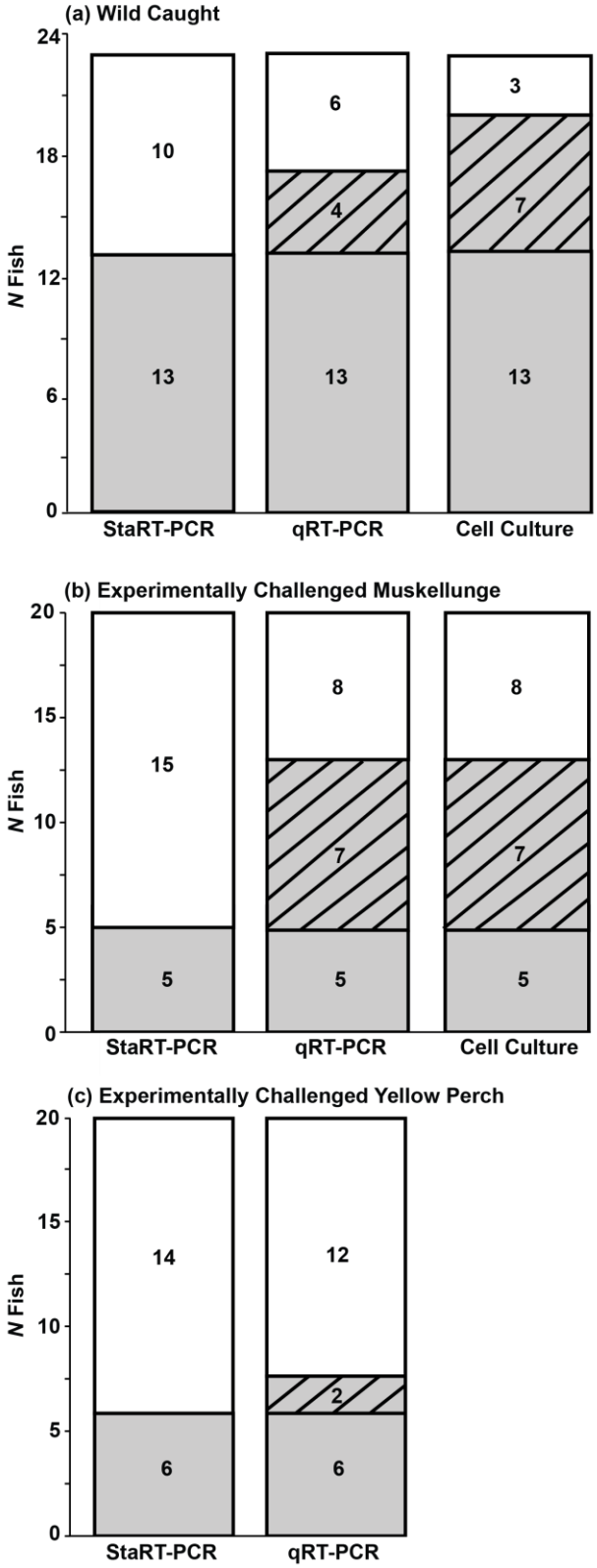

(a) Wild Caught
N Fish
24
18
12
6
0
10
13
6
4
13
3
7
13
StaRT-PCR
qRT-PCR
Cell Culture
(b) Experimentally Challenged Muskellunge
N Fish
20
15
10
5
0
15
5
8
7
5
8
7
5
StaRT-PCR
qRT-PCR
Cell Culture
(c) Experimentally Challenged Yellow Perch
N Fish
20
15
10
5
0
14
6
12
2
6
StaRT-PCR
qRT-PCR

Fig. 4-7

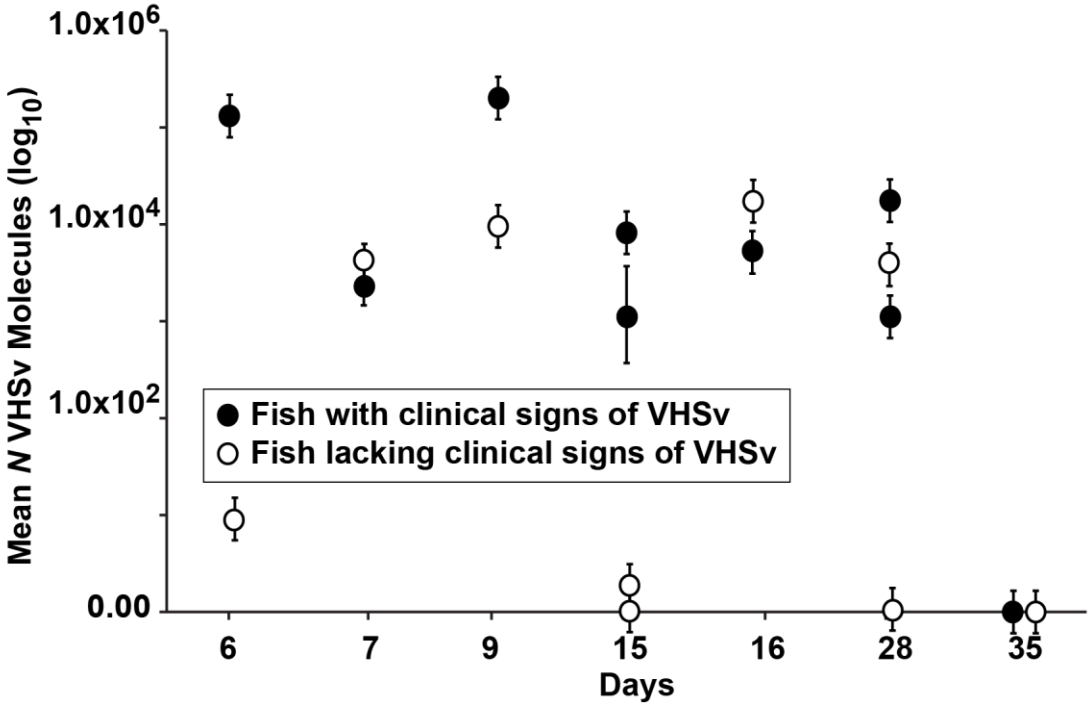
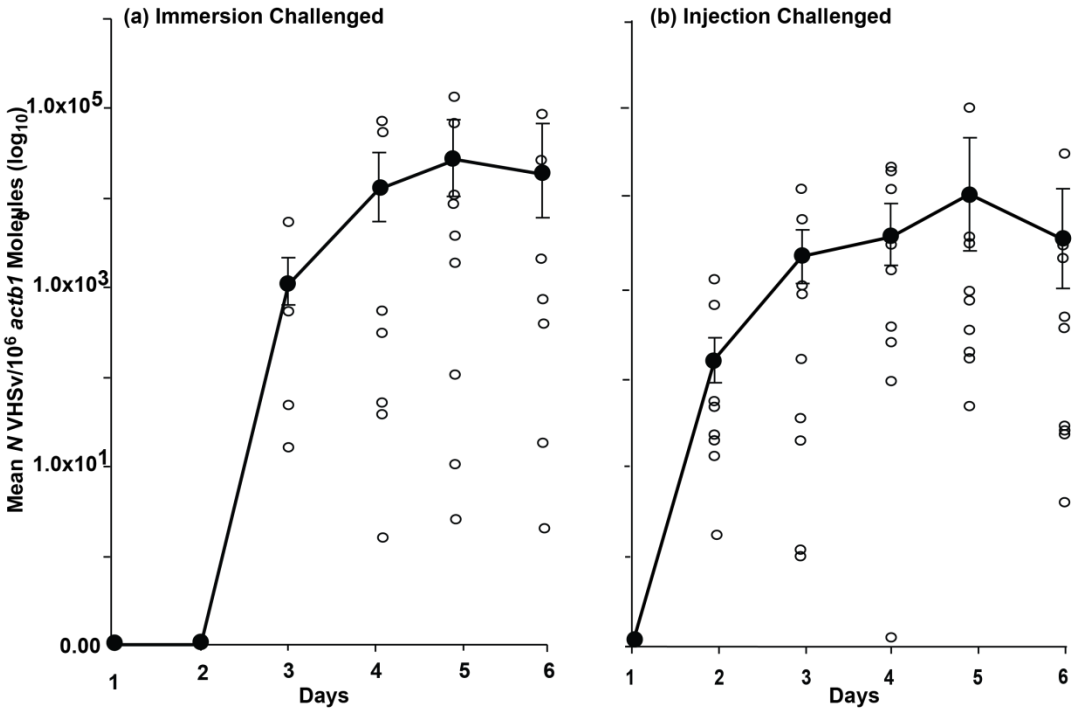

1.0x10^6
1.0x10^4
1.0x10^2
0.00
Mean N VHSv Molecules (log10)
Fish with clinical signs of VHSv
Fish lacking clinical signs of VHSv
6
7
9
15
16
28
35
Days

Fig. 4-8


(a) Immersion Challenged
(b) Injection Challenged
Mean N VHSv/10^6 actb1 Molecules (log10)
1.0x10^5
1.0x10^3
1.0x10^1
0.00
1
2
3
4
5
6
Days

Chapter 5

Accurate detection and quantification of the fish Viral Hemorrhagic Septicemia virus (VHSv) with a two-color fluorometric real-time PCR assay

Previously published as Pierce, L.R., Willey, J.C., Palsule, V.V., Yeo, J., Shepherd, B.S., Crawford, E.L., Stepien, C.A. (2013) Accurate detection and quantification of the fish Viral Hemorrhagic Septicemia virus (VHSv) with a two-color fluorometric real-time PCR assay. *PLoS ONE*

5.1 ABSTRACT: Viral Hemorrhagic Septicemia virus (VHSv) is one of the world's most serious fish pathogens, infecting >80 marine, freshwater, and estuarine fish species from Eurasia and North America. A novel and especially virulent strain – IVb – appeared in the Great Lakes in 2003, killed many game fish species in a series of outbreaks in subsequent years, and shut down interstate transport of baitfish. Cell culture is the diagnostic method approved by the USDA-APHIS, which takes a month or longer, lacks sensitivity, and does not quantify the amount of virus. We thus present a novel, easy, rapid, and highly sensitive real-time reverse quantitative PCR (qRT-PCR) assay that incorporates synthetic competitive template internal standards for quality control to circumvent false negative results. Results demonstrate high signal-to-analyte response (slope=1.00±0.02) and a linear dynamic range that spans seven orders of magnitude

($R^{2=}$0.99), ranging from 6 to 6,000,000 molecules. Infected fishes are found to harbor levels of virus that range to 1,200,000 VHSv molecules/10^6 *actb1* molecules with 1,000 being a rough cut-off for clinical signs of disease. This new assay is rapid, inexpensive, and has significantly greater accuracy than other published qRT-PCR tests and traditional cell culture diagnostics.

5.2 INTRODUCTION

Molecular diagnostic tools have facilitated the early detection, prevention, and spread of many important pathogens (Rao et al., 2006), led by the speed, sensitivity, and accuracy of Polymerase Chain Reaction (PCR)-based assays (Park et al., 2011). Their ability to diagnose targeted genetic sequences and quantify levels of infectious agents with hybridization probes has advanced screening technology for multiple human diseases, including influenza, hepatitis, and HIV (Coutlée et al., 1991; Ellis and Zambon, 2002). Use of these approaches to elucidate and characterize plant and animal pathogens likewise is growing at a rapid pace (Chai et al., 2013; Pasche et al., 2013).

Viral Hemorrhagic Septicemia virus (VHSv) causes one of the world's most serious finfish diseases, infecting >80 species across the Northern Hemisphere (Faisal et al., 2012), yet there remains a need for a fast, sensitive, accurate, and inexpensive diagnostic test. VHSv is a negative-sense, single stranded RNA *Novirhabdovirus* of ~ 12,000 nucleotides, with six open reading frames of 3'*N-P-M-G-Nv-L*'5 (Ammayappan and Vakharia, 2009). Infected fishes often swim erratically and have bulging eyes, distended abdomens, and extensive external/internal hemorrhaging (Winton and Einer-

Jensen, 2002). The virus survives for up to 13 days in the water (Hawley and Garver, 2008), and can be spread via ballast water, boating, equipment, and aquatic animals (e.g. birds, turtles, leeches, and amphipod crustaceans) (Meyers and Winton, 1995; Bain et al., 2010; Goodwin and Merry, 2011; Faisal and Winters, 2011). It is transmitted most readily during the spring spawning season through fish waste, reproductive fluids, and skin secretions (Meyers and Winton, 1995).

VHSv first was described in the late 1930s as "*Nierenschwellung*" in aquacultured rainbow trout (*Oncorhynchus mykiss*) from Europe (Shäperclaus, 1938). It now occurs across the Northern Hemisphere as four genetically and geographically distinct strains (I-IV) and substrains, whose evolutionary and biogeographic patterns recently were analyzed by Pierce and Stepien (2012). Strains I-III primarily occur in Europe, where they infect a wide variety of marine, estuarine, and freshwater fishes. Strain IV (now classified as IVa; (Elsayed et al., 2006)) first was discovered in 1988 from North American Pacific coastal fishes, including salmonids (Brunson et al., 1989; Hopper, 1989), and now also occurs in Japan (Takano et al., 2000). In 2000, another IV substrain (now designated as IVc per (Pierce and Stepien, 2012)) was discovered off the coast of New Brunswick, Canada, infecting the estuarine mummichog (*Fundulus heteroclitus*) and three-spined stickleback (*Gasterosteus aculeatus*) (Gagné et al., 2007). In 2003, a new and especially virulent substrain, IVb, was described from a moribund muskellunge (*Esox masquinongy*) in Lake St. Clair of the freshwater Laurentian Great Lakes system (Elsayed et al., 2006). Substrain IVb since has spread throughout all five of the Great Lakes, infecting >30 species, including many commercially and ecologically important fishes, such as muskellunge, drum (*Aplodinotus grunniens*), walleye (*Sander vitreus*),

yellow perch (*Perca flavescens*), and round goby (*Neogobius melanostomus*). Substrain IVb now contains at least 16 glycoprotein (*G*)-gene sequence variants (Thompson et al., 2011), whose rapid spread and diversification in a quasispecies mode suggest that this strain mutates rapidly and may be highly adaptable (see Pierce and Stepien (2012)).

To avoid outbreaks of the virus, the Aquatic Invasive Species Action Plan (2011) requires that aquaculture and baitfish vendors from U.S. states (Illinois, Indiana, Michigan, Minnesota, New York, Ohio, Pennsylvania, and Wisconsin) and Canadian provinces (Ontario and Quebec) have their fish products certified as VHSv-free prior to interstate or international transport. Cell culture is the VHSv diagnostic that is approved by the World Organization for Animal Health (Office of International des Epizooties – OIE, 2013), along with the joint Fish Health Section of the U.S. Fish and Wildlife Service and the American Fisheries Society (2010). The cell culture process takes a month or longer for cell growth, cell confluency, viral growth, and confirmation PCR. It moreover lacks the sensitivity to detect low viral concentrations in carrier fish, and results in false negative levels reported as 43-95% (Chico et al., 2006; López-Vázquez et al., 2006; Jonstrup et al., 2013).

Real-time quantitative reverse transcription (qRT)-PCR assays for detecting VHSv (Chico et al., 2006; López-Vázquez et al., 2006; Liu et al., 2008; Matejusova et al., 2008; Cutrín et al., 2009; Hope et al., 2010; Garver et al., 2011; Phelps et al., 2012; Jonstrup et al., 2013) likewise have substantially high false negative rates that ranged from 15-92% (Chico et al., 2006; López-Vázquez et al., 2006; Jonstrup et al., 2013). The high false negative rates in those assays may have resulted from unknown and/or unmonitored effects from interfering substances in the PCR or reverse transcription

reactions (rxn), which circumvented detection of the target gene (Huggett et al., 2008).

To avoid those issues, the present research describes and evaluates a new, accurate, fast, and highly reliable assay to diagnose and quantify VHSv. This assay uses Standardized Reverse-Transcriptase Polymerase Chain Reaction, i.e. StaRT-PCR, which is a form of competitive template RT-PCR that yields rapid, reproducible, standardized, and quantitative measurement of data for many genes simultaneously (Willey et al., 1998). StaRT-PCR uniquely incorporates internal standards (IS) in the rxn mixture to improve accuracy and circumvent false negative results. Our new assay is based on real-time PCR equipment that is readily available in most diagnostic laboratories, markedly improving on a previous version of our VHSv test (Pierce et al., 2013), which also used StaRT-PCR, but relied on less common capillary electrophoresis equipment (i.e. Agilent; Agilent Technologies, Santa Clara, CA). In the present study, results from both assays are evaluated by us to determine the presence or absence of VHSv and measure concentration of the virus from fish samples in targeted field and laboratory studies. We assay the VHSv nucleoprotein (*N*)-gene and the fish reference beta-actin 1 (*actb1*) gene, assessing amplification relative to known numbers of their respective competitive IS molecules. Our new approach uses sequence specific fluor-labeled hydrolysis probes that can be used on a variety of real-time PCR thermal cyclers, on which positive VHSv results are visible as two colors on the real-time PCR plot (see Fig. 5-1a; green=native template (NT), red=IS).

Results of our 2-color fluorometric assay are compared to those from our previously reported Agilent capillary electrophoresis test (Pierce et al., 2013), SYBR® green qRT-PCR, and cell culture (OIE, 2013), using the same biological samples. The

numbers of VHSv molecules are quantified from field-caught and laboratory-challenged VHSv-infected fish samples using the new 2-color fluorometric assay in comparison to the Agilent capillary electrophoresis test (Pierce et al., 2013).

5.3 MATERIALS AND METHODS

5.3.1 VHSv assay development

All primers and NT probes were matched to homologous sequences of the VHSv *N*-gene, based on all VHSv strains and substrain variants from NIH GenBank (http://www.ncbi.nlm.nih.gov/genbank/) and the literature, using Biosearch Technologies Real Time Design software (Novato, CA; http://www.biosearchtech.com/). The original muskellunge isolate MI03GL from the Great Lakes (GenBank Accession no. DQ427105) served as the reference for VHSv, and *actb1* mRNA from the yellow perch *Perca flavescens* (AY332493) was used as the fish reference gene sequence. Selection criteria included: product lengths that were <100 bp, with optimal melting temperatures of 65-68^{o}C for primers and 68-72^{o}C for probes. NT probes for the target and reference genes were labeled with FAM (fluorescein amidite dye).

The competitive template IS probes for the VHSv *N*-gene and the fish reference *actb1* gene each were constructed by altering 5-6 bp of the NT probe sequences, and were labeled with Quasar dye having 670 nm maximum absorbance (Biosearch Technologies). The IS probes were designed to: minimize cytosine and thymine (30%), maximize adenine (50%), avoid guanine at the 5' end, have lengths <24 bp, and have predicted

melting temperatures ± 0.02^{o}C of the NT probe. Synthetic NT and IS templates for VHSv and *actb1* were assembled by combining the forward primer, probe, and connecting sequence through the reverse primer (Table 5.1), and were synthesized by Life Technologies (Grant Island, NY; http://www.lifetechnologies.com/us/en/ home.html). The BLAST procedure (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was employed to verify that all primers, probes, and IS did not recognize other viral or fish DNA sequences.

To ensure that the probes did not bind to non-homologous template, their specificities were tested using synthetic templates for the VHSv *N*- and *actb1* genes. Both synthetic templates (NT and IS) were serially diluted 10-fold from 10^{-11} M to 10^{-15} M and tested with all probes in PCR amplification experiments, following the directions for "*Performing the VHSv Assay*", as detailed below. For example, the VHS *N*-gene IS synthetic template was evaluated with the VHS *N*-gene IS probe, as well as the VHS *N*-gene NT probe, and vice versa. The same was done for *actb1*. Cycle thresholds (C_t) from the homologous and non-homologous templates were compared at each dilution, and the non-homologous amplifications were calculated with formula $2^{(-\Delta Ct)}$ and multiplied by the known number of input copies. If the resulting numbers of molecules were >10% of the known input copy number, then the probe was re-designed, and the process was repeated.

After synthesis, the NT and IS for each gene were PCR amplified (Table 5.1) in 10 individual 10 µl rxns, containing 1 µl 750 nM of each primer, 0.5 U Go-TAQ polymerase (Promega, Madison, WI), 1 µl 10X $MgCl_2$ PCR buffer, 0.2 mM dNTPs, and RNAse-free water on a Rapid Cycler 2 (Idaho Technology, Inc., Salt Lake City, UT; www.biofiredx.com/). Rxns were run for 35 cycles of 5 sec at 94^{o}C, 10 sec at 58^{o}C, and 15 sec at 72^{o}C, with a slope of 9.9. To purify the NT and IS, all 10 replicate PCR

products per template were combined into a single tube, loaded onto individual 2% low melting pre-cast agarose gels from E-Gel iBase (Invitrogen, Grand Island, NY; www.invitrogen.com/), separated by electrophoresis, and visualized on a UV transilluminator. The NT and IS bands for the VHSv *N*- and *actb1* genes were harvested from their respective gel collection wells. Mean concentrations (ng/μl) of each were calculated from 1 μl of their purified products as measured with an Agilent 2100 Bioanalyzer in triplicate, and converted into molarities according to the formula (1):

(1) Molarity=[concentration (μg/μl)]/[molecular wt. (μg/μmole)]

To control for inter-sample and inter-experimental pipetting variation, a synthetic internal standard mixture (ISM) was created with the purified IS described above. To prepare the original stock ISM "A", we estimated the relative concentrations of the VHSv *N*-gene and *actb1* IS needed to achieve a 1:1 cDNA NT:IS relationship in a variety of samples (Table 5.2). Briefly, we mixed 10^{-10} M of the VHSv *N*-gene IS and 10^{-11} M of the *actb1* IS in an initial stock, labeled ISM "A". To measure various levels of gene expression, other ISM mixtures (ISM B-H) were constructed using 10-fold serial dilutions of the VHSv *N*-gene relative to a constant concentration of the *actb1* gene IS at 10^{-11} M (Table 5.2). Additional 10-fold dilutions of each ISM (A-H) stock then were made with 0.1 ng/μl yeast tRNA carrier (Invitrogen, Carlsbad, CA) to prevent adherence of negatively charged IS molecules to the tube or pipette tip surfaces (Table 5.2, rows 2-8).

An External Standardized Mixture – ESM (comprised of the synthesized NT and

IS for the VHSv *N*- and *actb1* genes) – was made to control for inter-lot and inter-experimental variation in probe fluorescence intensity, guard against inter-experimental variation in C_t selection, and normalize the probe (see equation (2), "Correction for variation in fluorescence among probes"). Stock ESM containing 10^{-11} M NT/10^{-11} M IS for the VHSv *N*- and *actb1* genes was diluted to a working concentration of 10^{-13} M NT/10^{-13} M IS and 10^{-14} M NT/10^{-14} M IS with yeast tRNA (Invitrogen).

5.3.2 Fish samples used to evaluate the VHSv assay

Spleen tissues from a variety of fish samples were used to test our assay for the presence and concentration of VHSv (and to compare our results to other assays, using the same samples). Fish were obtained, maintained, anesthetized, and sacrificed following the Institutional Animal Care and Use Committee (IACUC) approved protocols from the University of Toledo (#106419), Michigan State University (MSU; East Lansing, MI) (#AUF 07/07-123-00), and the U.S. Geological Survey's (USGS) Western Fisheries Research Center Challenge Facility (WFCCF; Seattle, WA) (#2008-17). Fish were euthanized with an overdose of 25 mg/ml tricaine methanesulfonate (MS-222; Argent Chemical Lab, Redmond, WA) and decapitated to ensure death. To remove any external viral particles, each fish was washed separately by submerging it 3X in double distilled H_2O. The surgical site (anus to operculum) was disinfected with 100% ethanol and betadine using sterile equipment. Spleen tissue was removed, placed into individual 1.5 mL eppendorf tubes, flash frozen in liquid nitrogen or stored in RNAlater (Qiagen), and kept at -80^{o}C until further processing. Gloves and all equipment were changed

between each fish to ensure sterile conditions. Specimens were disposed of following the respective approved biohazard protocols of the University of Toledo, MSU, and USGS.

Samples tested for VHSv included cDNA from 23 wild-caught Great Lakes fishes, including 10 infected and 13 negatives: two bluegill (*Lepomis macrochirus*), a brown bullhead (*Ameiurus nebulosus*), a freshwater drum, seven largemouth bass (*Micropterus salmoides*), a smallmouth bass (*Micropterus dolomieu*), and 11 lake herring (*Coregonus artedi*). We also tested 40 fish from VHSv laboratory challenge experiments, including 20 muskellunge (15 VHSv infected and 5 negative controls) from the MSU-Aquatic Animal Health Laboratory (AAHL), and 20 yellow perch (14 VHSv-infected and 6 negative controls) from USGS-WFRCCF.

A series of laboratory challenge experiments were conducted by MSU-AAHL on certified VHSv-free juvenile muskellunge (Rathburn National Fish Hatchery, Moravia, Iowa) under MSU IACUC protocols AUF 07/07-123-00. Muskellunge were challenged via water immersion for 90 min with VHSv-IVb (isolate MI03GL) at $4.0x10^3$ pfu/ml, and the negative controls with 1 ml sterile maintenance minimum essential media. Fish then were placed into clean VHSv-free water, and later randomly sacrificed at pre-determined intervals, as previously described.

We also analyzed RNA from a series of juvenile yellow perch laboratory challenge experiments, using six-month-old (VHSv-certified-free) Choptank broodstrain (Rosauer et al., 2011) from the University of Wisconsin-Milwaukee's Great Lakes WATER Institute (Milwaukee, WI), which were conducted at USGS- WFRCCF under their 2008-17 IACUC protocol. Perch were challenged either via intra-peritoneal injection of $1.0x10^5$ pfu/ml VHSv-IVb (strain MI03GL) or with immersion for two hours

in the same dosage, while control groups had a dose of minimum essential media. Fish were selected randomly in days 0-6 for euthanization with 240 mg/L MS-222 and 1.2g/L $NaHCO_3$. Dissection followed protocols described above.

5.3.3 Performing the VHSv assay

Spleen tissue (0.25-0.50 g) was ground using a sterile mortar and pestle under liquid nitrogen, and its RNA was extracted with the TriREAGENT® (Molecular Research Center, Inc., Cincinnati, OH) protocol. The RNA was re-suspended in 30 µl RNase-free water, quantified with a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA), and adjusted to a 1 µg RNA/µl concentration. DNA-*free* DNase Treatment and Removal Reagents (Ambion Life Technologies, Grand Island, NY) were used to eliminate any contaminating gDNA. The purified RNA was reverse-transcribed to cDNA with 1 µg RNA, 5X First Strand buffer, 10 mM dNTPs, 0.05 mM random hexamers, 25 U/µl RNasin, and 200 U/µl M-MLV in a 90 µl rxn volume at 94°C for 5 min, 37°C for 1 h, and 94°C for 5 min. The cDNA was stored at -20°C.

A set of PCR rxns was run per each cDNA sample to determine the appropriate concentrations of NT and IS for *actb1* to achieve a ratio of >1:10 and <10:1 of amplified products. Once the NT:IS products were in balance, the VHSv *N*- and *actb1* target genes were pre-amplified simultaneously to increase the signal (i.e. lower C_t) relative to non-specific background.

For each pre-amplification, a 10 µl volume of a master mixture was prepared with 5 µl TaqMan® Universal Master Mix II (without uracil N-glycosylase; Applied

Biosystems International (ABI), Grand Island, NY), 1 μl of 10X primer solution (final concentration: 75 nM) of the forward and reverse primers for the VHS *N* and *actb1* genes (mixed together), and RNAse-free water. Eight μl of this master mixture was dispensed into individual wells of 0.1 mL 96-well TempPlate® (USA Scientific, Inc.; www.usascientific.com/) containing 1 μl cDNA and 1 μl of the appropriate ISM concentration (Table 5.2). This was done in triplicate to allow calculation of the mean and standard error (S.E.) of the relative VHSv *N*-gene concentration/10^6 *actb1* molecules per fish sample. The plates then were sealed with a TempPlate® RT Optical Film and centrifuged for 2 min at 2000 rpm. PCR rxns were conducted on an ABI 7500 Fast using standard mode cycling conditions: 10 min at 95^{o}C, followed by 13 cycles of 15 sec at 95^{o}C and 1 min at 60^{o}C. For the Poisson distribution experiments, 25 pre-amplification cycles were used due to lower amount of starting template. Three no-template controls per rxn, located on separated areas on the plate, were used to control for possible contamination.

A second round of amplification was performed, in which each pre-amplified sample was diluted 1000-fold with TE buffer (10 mM Tris-Cl, 0.1 mM EDTA, pH 7.4); 2 μl of each diluted product was placed into each well of a new 0.1 mL 96-well TempPlate®, along with 18 μl of a master mixture containing 10 μl TaqMan® Universal Master Mix II (without uracil N-glycosylase), 2 μl of each 10X primer solution (final concentration: 750 nM), 2 μl of each NT and IS probe (final concentration: 200 nM), and RNAse-free water. This second amplification was conducted as described above, except run for 40 cycles. The number of VHSv molecules/10^6 *actb1* molecules was calculated using equations (2) and (3) below.

(2) Measured NT molecules for the VHSv *N*- and *actb 1* genes:

(a) Correction for variation in fluorescence among probes=ESM{[measured NT cycle threshold (C_t)]–[measured IS C_t]}

(b) Measured NT signal relative to the IS signal for each gene (ΔC_t)={[NT C_t]–[IS C_t]}–normalizing value

(c) Measured NT molecule copy number=[# input IS molecules from ISM]x[$2^{(-\Delta Ct)}$]

(3) Final number of molecules in the cDNA sample and the control for pipetting:

VHSv molecules/10^6 *actb1* molecules=[VHSv measured NT]/[*actb1* measured NT]x10^6

5.3.4 Specificity, true accuracy, and linearity

Our assay was tested for non-specific amplification using two human viruses (Encephalomyocarditis virus and Vesicular Stomatitis virus) and five fish viruses related to VHSv (Hirame rhabdovirus, Infectious Hematopoietic Necrosis virus, Infectious Pancreatic Necrosis virus, Spring Viremia of Carp virus, and Snakehead rhabdovirus). The Snakehead rhabdovirus is the nearest relative to VHSv, with 62% sequence similarity (Ammayappan and Vakharia, 2009; Pierce and Stepien, 2012). Twenty-five VHSv isolates were tested to evaluate amplification across a range of European, Asian, and North American variants (Table 5.3), encompassing all four strains. All samples were assayed in triplicate.

To measure true accuracy – the agreement between a measurement and its known value (Shabir, 2003) – the relationship between the observed versus expected numbers of VHSv *N*-gene and *actb1* molecules based on Poisson analysis was determined (Vogelstein and Kinzler, 1999). Ten replicates were amplified for the VHSv *N*- and *actb1* genes over a series of limiting PCR dilutions, which were predicted to contain 40, 20, 10, 7, 6, 5, 4, 2, 1, 0.7, 0.4, and 0.1 molecules. Linear regression analysis was performed in the R statistical software suite v2.15.2 (R Core Development Team, 2012). A χ^2 test (in Microsoft Excel) compared the number of molecules measured with the 2-color fluorometric assay versus those from the Agilent 2100 Bioanalyzer, at the same dilutions.

Linearity was measured over two series of dilution experiments to: 1) determine the maximum and minimum ratio of NT to IS that produced reproducible results, and 2) verify that our test followed a linear trend in calculating the expected number of molecules per dilution. The first dilution set was made by mixing a constant amount of synthetic NT with decreasing amounts of IS to generate dilutions of: 1:1 ($6x10^4$ molecules), 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:12, 1:14, 1:16, 1:18, and 1:20 ($3x10^3$ molecules) for both genes. Identical procedures were performed by holding the IS constant, while varying the NT. The second dilution series evaluated linearity for the VHSv *N*- and *actb1* genes using 10-fold serial dilutions of the ESM at a 1:1 ratio, with dilutions of $6x10^6$, $6x10^5$, $6x10^4$, $6x10^3$, $6x10^2$, 60, 6, and 0.6 molecules. Regression analyses were conducted to determine correlation (R^2), slope (linearity), and relation to a linear trend (F) among the various dilutions of NT:IS and IS:NT for each gene. Imprecision was reported as the coefficient of variation (CV), calculated as the standard deviation divided by the mean of triplicate measurements at each dilution (reported in %)

(in Microsoft Excel) (Ogino et al., 2006). In addition, S.E. was calculated for each sample. For these linearity experiments, PCR was done as specified above in "Performing the VHSv Assay", but substituting the cDNA and ISM with either 2 μl of the ESM (from dilution 1) or a concentration of 1:1 NT/IS (dilution 2). Each dilution was run in triplicate, with a negative/no template control for each run.

5.3.5 VHSv detection comparisons of our assay to others

Results from the new 2-color fluorometric test are compared to those from our prior Agilent capillary electrophoresis assay (Pierce et al., 2013), conventional SYBR® green qRT-PCR, and cell culture to evaluate their relative abilities to detect VHSv in 63 fish samples (see "*Fish Samples used to Evaluate the VHSv Assay*"). All samples were analyzed in triplicate and all runs had positive and negative controls. Each PCR rxn included a known cell culture positive, a negative VHSv cDNA, and a reagent negative control (nuclease-free H_2O). PCR products were visualized on 1% agarose gels to confirm positive/negative results. The amount of yellow perch fish tissue available from the USGS laboratory challenge experiments precluded analysis with cell culture. χ^2 tests (Sokal and Rohlf, 1995) were used to compare the results among the approaches.

SYBR® green qRT-PCR experiments used a Mastercycler Realplex Thermocycler (Eppendorf, Inc., Westbury, NY) in 25 μl rxns, containing 0.05 μg of each primer (the same primers used for the Agilent capillary electrophoresis assay (Pierce et al., 2013)), 2 μl cDNA product, 10 μl SsoFast SYBR® green mix, and RNase-free water. Amplifications were run on a Mastercycler Realplex Thermocycler (Eppendorf, Inc.,

Westbury, NY), with initial denaturation of 5 min at 95°C, followed by 40 cycles of 30 sec at 95°C, and 1 min at 60°C.

Cell culture was performed at MSU's AAHL by M. Faisal and R. Kim, following standard OIE (2013) procedures. If results were positive, RNA was extracted from infected cells as described above, reverse transcribed with Affinity Script Multiple Temperature Reverse Transcriptase PCR (Stratagene, La Jolla, CA), and amplified following OIE (2013).

5.3.6 VHSv quantification using our assay

Positive samples were quantified using our new 2-color fluorometric real-time PCR assay and compared to our earlier results from the Agilent capillary electrophoresis procedure (Pierce et al., 2013) from the 63 test fish, with linear regression in R and an *F*-test (Sokal and Rohlf, 1995). Numbers of VHSv/10^6 *actb1* molecules were measured in triplicate, from which means and S.E. were calculated. Relative numbers of VHSv molecules were compared between laboratory challenged muskellunge showing clinical signs of infection (e.g. external hemorrhages; N=9) versus those without signs (N=9). A χ^2 test (Microsoft Excel) was used to determine if a threshold number of VHSv molecules characterized the appearance of the clinical signs. Due to limited sample size, a power analysis (G*Power2; (Erdfelder et al., 1996)) was used to estimate the number of fish needed to achieve 95% confidence, with an effect size of 0.50 (Cohen, 1992).

5.4 RESULTS

5.4.1 Performance of our 2-color fluorometric assay for VHSv

Our test results are negative for all other viruses (i.e. did not result in amplification; Table 5.3), including human viruses (Encephalomyocarditis virus and Vesicular Stomatitis virus) and fish viruses that are related to VHSv (i.e. Hirame rhabdovirus, Infectious Hematopoietic Necrosis virus, Infectious Pancreatic Necrosis virus, Spring Viremia of Carp virus, and Snakehead rhabdovirus). All four VHSv strains (I–IV) and all substrains evaluated (I, Ia, II, III, IVa, IVb, and IVc) yield positive amplification results with our test (Fig. 5-1a; Table 5.3). Thus this new assay is specific for VHSv.

Amplification results for the VHSv *N*-gene (Fig. 5-2a) are 100% (10/10 times) for dilutions of 5-40 VHSv molecules, 90% (9/10 times) for 4 molecules, 80% (8/10) for 2 molecules, 60% (6/10) for a single molecule, 30% (3/10) for 0.7 molecules, 20% (2/10) for 0.4 molecules, and 10% (1/10) for 0.1 molecules ($R^{2=}$0.98, F=541.50, df=1, 10, p<0.001). Values for amplification of the fish *actb1* gene are similar (Fig. 5-2b), yielding 100% positives for 4-40 molecules (10/10), 70% at 2 molecules (7/10), 40% for a single molecule and for 0.7 molecules (4/10), 20% for 0.4 molecules (2/10), and 10% at 0.2 molecules (1/10) ($R^{2=}$0.97, F=283.60, df=1, 10, p<0.001). Results indicate that the numbers of ISM molecules measured by our assay match those from the bioanalyzer for the VHSv *N*-gene ($\chi^{2=}$0.18, df=11, NS) and the *actb1* gene ($\chi^{2=}$0.23, df=11, NS).

The relationship between the amount of PCR product remains linear when the concentration of NT is held constant and the IS is varied for both the VHSv *N*-gene (Fig. 5-3a: $R^{2=}0.99$, $F=1363.00$, df=1, 13, $p<0.001$) and the *actb1* gene (Fig. 5-3b: $R^{2=}0.99$, $F=1283.00$, df=1, 13, $p<0.001$). Figure 5-1b depicts the results that illustrate this relationship. The mean calculated CV is 5% for the VHSv *N*-gene over an NT:IS dilution range of 1:1-1:10 (yielding $6.0x10^4 \pm 0.0x10^0$ to $4.6x10^3 \pm 2.8x10^2$ molecules). The CV likewise is 5% for the *actb1* gene (yielding $6.0x10^4 \pm 0.00x10^0$ to $7.5x10^3 \pm 4.3x10^2$ molecules). At dilutions beyond 1:10, the CV increases to 7% for both the VHSv *N*-gene (yielding up to $1.5x10^3 \pm 1.1x10^2$ molecules) and the *actb1* gene (yielding up to $3.0x10^3 \pm 1.2x10^2$ molecules) when the NT is held constant.

Analogous results are obtained when the IS is held constant and the NT is varied for the VHSv *N*-gene (Fig. 5-1c and Fig. 5-3c: $R^{2=}0.99$, $F=5124.00$, df=1, 13, $p<0.001$) and the *actb1* gene (Fig. 5-3d: $R^{2=}0.99$, $F=2434.00$, df=1, 13, $p<0.001$). The mean CV for the IS:NT dilution range of 1:1-1:10 is 5% for the VHSv *N*-gene (yielding $6.0x10^4 \pm 0.0x10^0$ to $5.9x10^3 \pm 2.2x10^3$ molecules) and 3% for the *actb1* gene (yielding $6.0x10^4 \pm 0.0x10^0$ to $4.0x10^3 \pm 4.3x10^1$ molecules). At dilutions beyond 1:10, the CV increases to 7% for the VHSv *N*-gene (yielding up to $2.8x10^3 \pm 1.8x10^2$ molecules) and 6% for the *actb1* gene (yielding up to $1.4x10^3 \pm 6.2x10^1$ molecules) when the IS is held constant. Based on these findings, our quantifications are conducted in the range of 1:10 to 10:1 NT:IS to maximize accuracy.

Numbers of VHSv molecules show a linear relationship over seven orders of magnitude (serial dilutions of $6x10^6$ to $6x10^0$ molecules) when the NT:IS is 1:1 (Fig. 5-4a), with a slope of 1.00 ($R^{2=}0.99$, $F=9404.00$, df=1, 5, $p<0.001$). Figure 5-1d illustrates

this relationship, in which NT and IS increase by ≤3.2 C_t for each 10-fold serial dilution of the ESM. The mean CV for VHSv is estimated at 7% for samples of $6x10^6$ to $6x10^1$ molecules (measured as $6.5x10^6 \pm 5.2x10^5$ to $7.9x10^1 \pm 2.0x10^0$ molecules), and 9% when the range is extended to $6x10^0$ molecules (measured as $6.0x10^0 \pm 1.0x10^0$ molecules). Results for *actb1* have a similar trend (Fig. 5-4b), with a slope of 1.04 ($R^{2=}$0.99, F=1347.00, df=1, 5, p<0.001), a mean CV of 7% for $6x10^6$ to $6x10^1$ molecules (measured as $6.6x10^6 \pm 2.1x10^5$ to $7.8x10^1 \pm 8.0x10^0$ molecules), and 10% when the range is extended to $6x10^0$ molecules (measured as $3.0x10^0 \pm 0.4x10^0$ molecules). Stochastic sampling likely contributes to increased CV and S.E. in the measurements for 6 molecules.

5.4.2 VHSv detection and quantification comparison among methods

Results reveal that our present 2-color fluorometric real-time PCR assay and previous results from the Agilent capillary electrophoresis-based approach (Pierce et al., 2013) both discriminate identical positives and negatives (i.e. they have the same accuracy; Fig. 5-5; $\chi^{2=}$0.00, df=1, NS), and are free of false negatives (Fig. 5-5). In contrast, the cell culture results have 56% false negative error (Fig. 5-5a: $\chi^{2=}$9.36, df=1, p=0.002) and SYBR® green yields 33-44% false negative error (Fig. 5-5a,b: $\chi^{2=}$5.37-5.67, df=1, p=0.02). All positives detected by SYBR® green qRT-PCR and cell culture also are positive with both of our StaRT-PCR methods (2-color fluorometric real time and capillary electrophoresis). The false negative range for SYBR® green qRT-PCR is $1.0x10^0$-$1.6x10^2$ VHSv/10^6 *actb1* molecules (=$0.6x10^0$-$2.5x10^2$ VHSv molecules, as

quantified by our 2-color fluorometric method) and $1.0x10^{0}$-$2.2x10^{3}$ VHSv/10^{6} *actb1* molecules (=$0.6x10^{0}$-$6.1x10^{3}$ VHSv molecules, as quantified by our 2-color fluorometric method) for cell culture. True negatives (including experimental controls) are negative with all assays; i.e. we find no false positives and no contamination.

Numbers of VHSv molecules/10^{6} *actb1* molecules measured from the spleen tissues of positive fish are higher in the new assay, ranging to $1.21x10^{6}$ VHSv molecules/10^{6} *actb1* (=$1.90x10^{4}$ VHSv molecules) than for the Agilent capillary-based test, which range to $8.4x10^{5}$ VHSv molecules/10^{6} *actb1* (=$2.7x10^{3}$ VHSv molecules). However, both sets of values have a direct linear relationship (Fig. 5-6: R^{2}=0.91, df=1, 38, F=396.40, p<0.001, t=1.42, df=78, NS). Muskellunge exhibiting clinical signs of infection contain a greater mean number of viral molecules ($1.4x10^{5}$±$6.5x10^{3}$ VHSv/10^{6} *actb1* molecules=$6.9x10^{4}$±$6.9x10^{3}$ VHSv molecules) than those without ($1.2x10^{4}$±$1.7x10^{3}$ VHSv/10^{6} *actb1* molecules=$1.5x10^{3}$±$1.6x10^{2}$ VHSv molecules). The estimated threshold at which those individuals display clinical signs of infection is ~$1x10^{3}$ VHSv/10^{6} *actb1* molecules (=$3.6x10^{2}$ VHSv molecules) using our assay. Our sample sizes are not sufficient to further evaluate the relationship between this threshold number of molecules and clinical diagnosis ($\chi^{2=}$0.09, df=1, NS). Power analysis estimates that 52 fish samples (26 with and 26 without clinical signs) would be needed to verify this finding.

5.5 DISCUSSION

Disease diagnostic laboratories depend on rapid, sensitive, and accurate detection methods, which are easy to employ and cost-effective. Cell culture is the VHSv diagnostic approved by the World Organization of Animal Health (OIE, 2013), which takes up to a month to perform in clinical settings and often results in substantial false negatives – as revealed here and by other studies (Chico et al., 2006; López-Vázquez et al., 2006; Jonstrup et al., 2013). Compared with traditional cell culture, all PCR-based assays (Hope et al., 2010) – including the present one – show enhanced ability to detect VHSv since they amplify both the negative-strand non-replicating genomic RNA and the positive-strand replicating mRNA transcripts. Amplification of both transcripts may be beneficial since positive results may denote new spread of VHSv or latent cases in the geographic region where the samples are taken. This can aid in diagnosis of viral infections.

The present assay detects and quantifies VHSv-IVb in fishes from the Great Lakes using primers and probes that are homologous to the *N*-gene sequence of the widespread IVb isolate MI03GL, and matches conserved sequence regions among all VHSv strains and substrains. Results demonstrate cross-reaction with all other VHSv strains and substrains tested. Other human and fish viruses do not amplify. Thus, our assay is VHSv-specific and detects all of its known variants.

Other PCR tests developed for VHSv by Chico et al. (2006), López-Vázquez et al. (2006), Liu et al. (2008), Matejusova et al. (2008), Cutrín et al. (2009), Hope et al. (2010), Garver et al. (2011), and Jonstrup et al. (2013), culminated in high numbers of

false negatives, analogous to the SYBR® green test evaluated here (33-44% false negatives). Notably, 15-90% false negatives were reported by López-Vázquez et al. (2006), 25-92% by Chico et al. (2006), and values to 42% by Jonstrup et al. (2013) for their respective approaches. Unlike those other real-time PCR tests for VHSv (Chico et al., 2006; López-Vázquez et al., 2006; Liu et al., 2008; Matejusova et al., 2008; Cutrín et al., 2009; Hope et al., 2010; Garver et al., 2011; Phelps et al., 2012; Jonstrup et al., 2013), our method incorporates intrinsic quality control standards (IS) to circumvent false negative results.

Specifically, exogenous (IS) and endogenous controls (the commonly used reference gene *actb1*) facilitate optimal detection of true positives and act to normalize the quantification of viral molecules. Use of IS is recommended by the International Organization for Standardization (2005), the U.S. Environmental Protection Agency (2004), and the U.S. Food and Drug Administration (2010). Tests for Hepatitis C virus (Gelderblom et al., 2006) and Human Immunodeficiency virus (Swanson et al., 2006; Schumacher et al., 2007) already have implemented IS in their assays.

Our assay is sensitive, follows a linear relationship with increasing viral concentration, and is highly reproducible. It detects down to five VHSv *N*-gene molecules with 100% accuracy, based on Poisson distribution. Other real-time PCR assays for VHSv had much higher detection thresholds. Notably, Liu et al.'s (2008) test required ≥140 viral copies of VHSv, and assays by Hope et al. (2010) and Garver et al. (2011) needed ≥100 viral copies. Our results are consistent for samples containing six to 6,000,000 VHSv molecules. Stochastic variation is evident only at extremely low dilutions (<five molecules). Results confirm reliability from concentrations of 1:1 to 1:20

NT:IS, with some slight increase in CV at dilutions >1:10. We thus recommend adjusting the relative concentrations of NT:IS to maximize accuracy, following recommendations in the Materials and Methods section "*Performing the VHSv Assay*". All quantification values reported here fall within this 1:10 range, which allows us to distinguish a ~1.25-fold C_t difference. Our assay also should work well with highly degraded samples (e.g. dead fish in the field), as described for human cancer samples using this type of approach by some of our team members (Yeo et al., unpublished data).

This 2-color fluorometric real-time assay is highly accurate and is free of the size separation steps required for our previously-published Agilent capillary electrophoresis approach (Pierce et al., 2013). Here we determine higher numbers of VHSv molecules for the same fish samples, due to the re-design of primers and use of fluorescent-labeled probes. Results from both methods have a linear relationship and are readily cross-calibrated.

Laboratory challenged muskellunge showing clinical signs of infection contain a greater mean number of viral molecules than those without. It is estimated that ~$1x10^3$ VHSv/10^6 *actb1* molecules (=$3.6x10^2$ VHSv molecules) appears to mark a clinical threshold for signs of the disease. However, exhibition of clinical signs at this biomarker could be species-specific, and may differ between fish in the laboratory and those in the field. Further experiments are warranted to validate this assumption.

5.5.1 Conclusions

Our assay is highly sensitive and accurate, free of false negatives, and reliably

quantifies a wide range of VHSv in fish tissue samples. Other PCR-based methods and cell culture techniques had high proportions of false negatives since they lacked intrinsic quality control, which could lead to spread of the virus. This new test will aid rapid, accurate, and low-cost diagnosis of VHSv. It has wide applicability across the geographic range of the virus, and should be highly successful in elucidating new occurrences and circumventing spread.

5.6 ACKNOWLEDGEMENTS

We thank T. Blomquist and L. Stanoszek for assistance, and P. Uzmann and M. Gray for logistic support. I. Bandín of Universidad de Santiago de Compostela (La Coruña, Spain), P. Bowser of Cornell University (Ithaca, NY), T. Gadd of the Finnish Food Safety Authority Evira (Helsinki, Finland), K. Garver of Fisheries and Oceans (Ottawa, Canada), F. Goetz of NOAA Northwest Fisheries Science Center Manchester Research Station (Port Orchard, WA), D. Leaman of the University of Toledo (Toledo, OH), and G. Kurath and J. Winton of USGS (Seattle, WA) provided VHSv samples. The views contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Government. Mention of trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the USDA and does not imply its approval to the exclusion of other products that may be suitable.

Table 5.1

Sequences and PCR parameters for our 2-color fluorometric VHSv assay. Primers, probes, internal standards (IS), and synthetic templates are specified. F=forward primer, R=reverse primer, NT=native template. *Italics*=modified nucleotides in NT probe.

Primers, Probes, and IS	Nucleotide position	Sequence (5'-3')	T_m (°C)	GC%	Product length (bp)
N-gene:					
pVHS4b F2	316-336	GCC GGA ATC CTT ATG CCG ATG	68.0	57.1	74
pVHS4b R2	367-389	CCC TTG ACG ATG TCC ATG AGG TT	67.0	52.0	74
Housekeeping Gene:					
p*actb1* F4	1075-1095	CCC ACC AGA GCG TAA ATA CTC	65.5	52.0	92
p*actb1* R4	1145-1165	CTC CTG CTT GCT GAT CCA CAT	65.6	52.0	92
Probes:					
pVHS 4b 1NT	342-359	ACT GGC CCA GAC TGT CAA	68.5	55.6	18
VHS4b IS	342-358	*TG*T GGC C*GA* G*TC A*GT C*C*	70.3	64.7	17
p*actb1*_4 NT	1098-1116	TCT GGA TCG GAG GCT CCA T	71.8	57.9	19
p*actb1*_4 IS	1098-1117	*ACT CGA TTG* GAG G*GT CCG AC*	71.8	60.0	20
Synthetic Templates:					
VHS 4b NT	304-416	ACT GGC ATC GAG GCC GGA ATC CTT ATG CCG ATG AAG GAA CTG GCC CAG ACT GTC AAC GCC GAC AAC CTC ATG GAC ATC GTC AAG GGG GCC CTG ATG ACG TGT TCC CTT CTG AC	82.2	56.6	113
VHS4b IS2	304-416	ACT GGC ATC GAG GCC GGA ATC CTT ATG CCG ATG AAG GA*T GTG* GCC *GAG TCA* GTC *C*AC GCC GAC AAC CTC ATG GAC ATC GTC AAG GGG GCC CTG ATG ACG TGT TCC CTT TCG AC	82.6	58.4	113
p*actb1*_4 NT	1075-1095	CCC ACC AGA GCG TAA ATA CTC TGT CTG GAT CGG AGG CTC CAT CCT GGC CTC TCT GTC CAC CTT CCA GCA GAT GTG GAT CAG CAA GCA GGA G	85.3	57.1	91
p*actb1*_4 IS	1145-1165	CCC ACC AGA GCG TAA ATA CTC TG*A* CT*C* GAT *T*GG AGG *G*TC C*GA C*CC TGG CCT CTC TGT CCA CCT TCC AGC AGA TGT GGA TCA GCA AGC AGG AG	85.5	57.6	92

Table 5.2

Concentrations for the 2-color fluorometric VHSv assay. Dilution mixtures (A-H) used for the Internal Standards Mixture (ISM) *actb1*/VHSv are given in units of 10^x M.

A	B	C	D	E	F	G	H
-11/-10	-11/-11	-11/-12	-11/-13	-11/-14	-11/-15	-11/-16	-11/-17
-12/-11	-12/-12	-12/-13	-12/-14	-12/-15	-12/-16	-12/-17	
-13/-12	-13/-13	-13/-14	-13/-15	-13/-16	-13/-17		
-14/-13	-14/-14	-14/-15	-14/-16	-14/-17			
-15/-14	-15/-15	-15/-16	-15/-17				
-16/-15	-16/-16	-16/-17					
-17/-16	-17/-17						

Table 5.3

Viral isolates screened using StaRT-PCR. -=negative result (no amplification), +=positive result.

Type	Isolate	Result
Human:		
Encephalomyocarditis virus		-
Vesicular Stomatitis virus		-
Fish:		
Hirame rhabdovirus[a]		-
Infectious Hematopoietic Necrosis virus (strain 220-90)[a]		-
Infectious Pancreatic Necrosis virus		-
Snakehead rhabdovirus[a]		-
Spring Viremia of Carp virus[a]		-
VHSv:		
I	DK-F1[b]	+
Ia	FR0771[b]	+
Ia	JP96KRRV9601[b]	+
II	FI-ka663-06[c]	+
III	SM2897[d]	+
III	SC2645[d]	+
III	GH35[d]	+
III	GH 44[d]	+
IVa	Bogachiel[b]	+
IVa	Cod'91[b]	+
IVa	Elliott Bay[b]	+
IVa	JP96Obama[b]	+
IVa	Makah[b]	+
IVa	Orcas[b]	+
IVb	MI03GL[a,b]	+
IVb	vcG002[a]	+
IVb	vcG003[a]	+
IVb	vcG004[a]	+
IVb	vcG005[a]	+
IVb	vcG006[a]	+
IVb	vcG007[a]	+
IVb	vcG008[a]	+
IVb	vcG009[a]	+
IVb	vcG010[a]	+
IVc	CA-NB00-02[e]	+

Isolates obtained from:
[a]Western Fisheries Research Center, USGS, Seattle, WA, USA
[b]Cornell University College of Veterinary Medicine, Ithaca, NY, USA
[c]Finnish Food Safety Authority Evira, Finland
[d]Universidad de Santiago de Compostela, Spain
[e]Fisheres and Oceans Canada, Pacific Biological Station, BC, Canada

Fig. 5-1 Real-time PCR amplification plots for various experiments. ABI 7500 real-time PCR results for (a) a true VHSv positive fish sample, (b) the relationship between the VHSv Native Template (NT) and Internal Standard (IS) with the NT held constant and the IS varied, (c) the relationship between VHSv NT and IS with the IS held constant and the NT varied, and (d) the relationship between VHSv NT:IS with concentrations held constant for dilutions of 1:1-1:20. Green=NT, red=IS, y=fluorescence of the reporter dye minus the baseline (Δ fluorescence), x= cycle threshold.

Fig. 5-2 True accuracy of the 2-color fluorometric test. Results are based on % positives from 10 separate runs of 12 dilutions using a known Internal Standard Mixture (ISM). Dilutions are: 40, 20, 10, 7, 6, 5, 4, 2, 1, 0.7, 0.4, and 0.1 molecules. The 2-color fluorometric test yields 100% positives for (a) ≥5 molecules of VHSv and (b) ≥4 molecules for *actb1*.

Fig. 5-3 Relationship between the number of observed and expected molecules for NT:IS concentrations of 1:1-1:20. The concentration of Native Template (NT) is held constant and the Internal Standard (IS) is varied for dilutions of: 1:1 ($6x10^4$ molecules), 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:12, 1:14, 1:16, 1:18, and 1:20 ($3x10^3$molecules). The 2-color fluorometric assay yields a linear relationship for (a) VHSv ($R^{2=}$0.99, F=1514.00, df=1, 13, p<0.001) with a mean CV of 5% for dilutions 1:1-1:10 and 7% for concentrations down to 1:20, and for (b) *actb1* ($R^{2=}$0.99, F=1283.00, df=1, 13, p<0.001), CV=5% and 7%. The same linear pattern is observed when the IS was held constant and NT varied for (c) VHSv ($R^{2=}$0.99, F=5124.00, df=1, 13, p<0.001), CV=5% for 1:1–1:10

and 7% for dilutions down to 1:20, and (d) *actb1* ($R^{2=}$0.99, F=2434.00, df=1, 13, p<0.001), CV=3% and 6%. Error bars=standard error of results for triplicate runs. Dotted line=partition of dilutions from 1:1-1:10 (right) and 1:12-1:20 (left).

Fig. 5-4 Relationship between the numbers of observed versus expected molecules when NT:IS concentrations are 1:1. Results are based on dilutions of the Native Template (NT) and Internal Standard (IS) of $6x10^6$, $6x10^5$, $6x10^4$, $6x10^3$, $6x10^2$, 60, 6, and 0.6 molecules for VHSv and *actb1*. The 2-color fluorometric assay yields a linear relationship for (a) VHSv over seven orders of magnitude (from $6x10^6$ to $6x10^0$ VHSv molecules) with a slope of 1.00 ($R^{2=}$0.99, F=9404.00, df=1, 5, p<0.001), and mean CV of 9%. A linear trend also is obtained for (b) *actb1* ($R^{2=}$0.99, F=1347.00, df=1, 5, p<0.001). Slope=1.04, mean CV=10%. Error bars=standard error of triplicate runs.

Fig. 5-5 Relative numbers of VHSv positives and negatives from our 2-color fluorometric and capillary electrophoresis StaRT-PCR assays, which indicates identical numbers of positives and negatives. Compared to these tests, for 43 fishes (25 positives, 18 negatives (including 5 laboratory controls)), (a) SYBR® green has 44% false negative error ($\chi^{2=}$5.67, df=1, p=0.02), and cell culture has a 56% error ($\chi^{2=}$9.36, df=1, p=0.002). For 63 fish samples (39 positives, 24 negatives (including 11 laboratory controls)), (b) SYBR® green qRT-PCR has 33% false negative error ($\chi^{2=}$5.37, df=1, p=0.02), whereas the 2-color fluorometric and capillary electrophoresis tests show zero false negatives.

Fig. 5-6 Mean log numbers of VHSv molecules/10^6 *actb1* molecules from our new 2-color fluorometric assay versus the prior Agilent capillary electrophoresis approach. Results show a linear relationship between the two tests (R^2=0.91, df=1, 38, F=396.40, p<0.001) and do not significantly differ (t=1.42, df=78, NS). Arrow=Estimated threshold concentration of VHSv for fish with clinical signs of infection using our new assay, from a χ^2 test of nine symptomatic fish ($1x10^3$ VHSv molecules/10^6 *actb1* molecules=$3.6x10^2$ VHSv molecules). Triangle=false negative range for SYBR® green qRT-PCR ($1.0x10^0$-$1.6x10^2$ VHSv/10^6 *actb1* molecules=$0.6x10^0$-$2.5x10^2$ VHSv molecules). Square=false negative range for cell culture ($1.0x10^0$-$2.2x10^3$ VHSv/10^6 *actb1* molecules=$0.6x10^0$-$6.1x10^3$ VHSv molecules.

Fig. 5-1

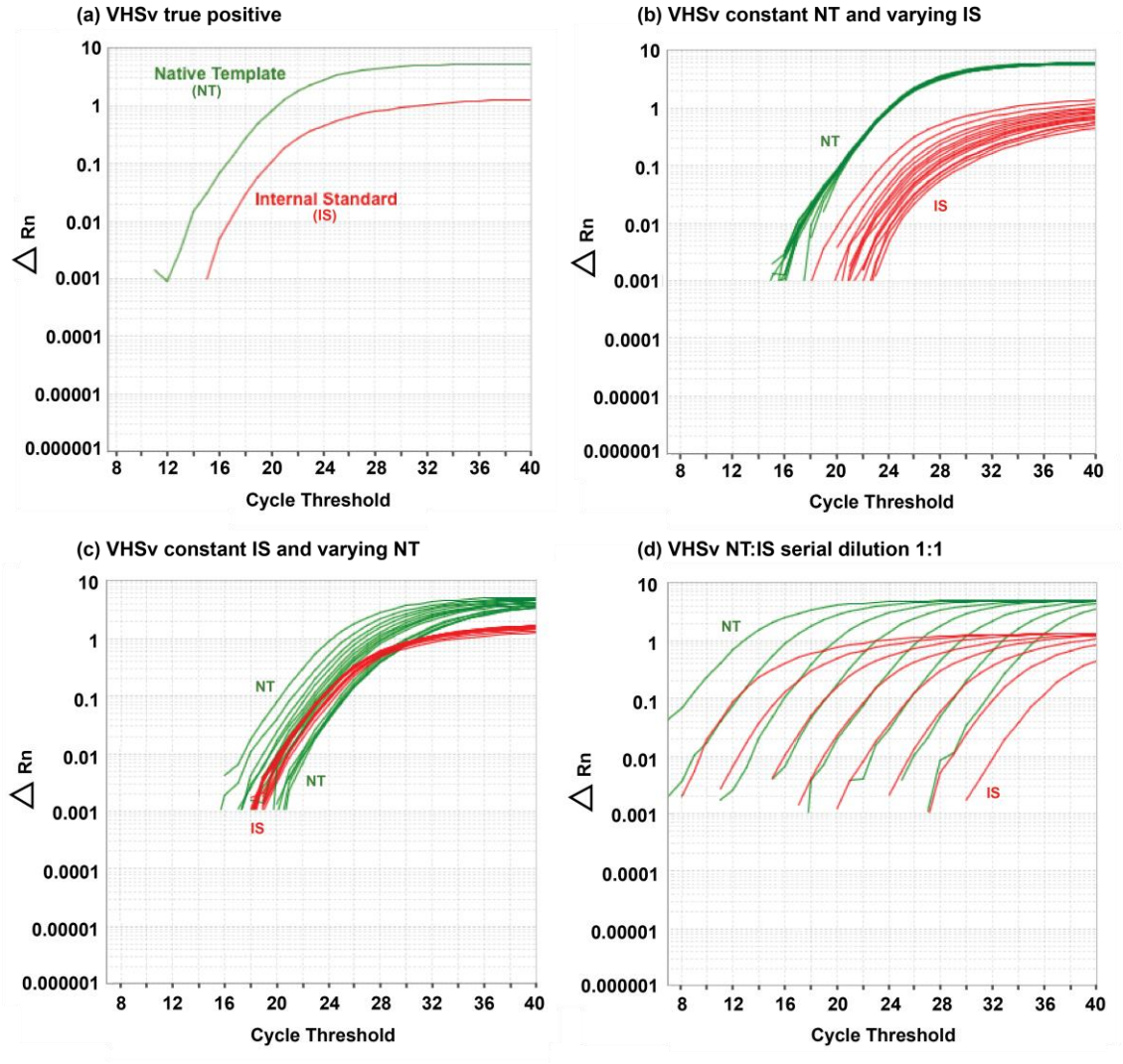

(a) VHSv true positive
Native Template
(NT)
Internal Standard
(IS)
ΔRn
Cycle Threshold
(b) VHSv constant NT and varying IS
NT
IS
ΔRn
Cycle Threshold
(c) VHSv constant IS and varying NT
NT
NT
IS
ΔRn
Cycle Threshold
(d) VHSv NT:IS serial dilution 1:1
NT
IS
ΔRn
Cycle Threshold

Fig. 5-2

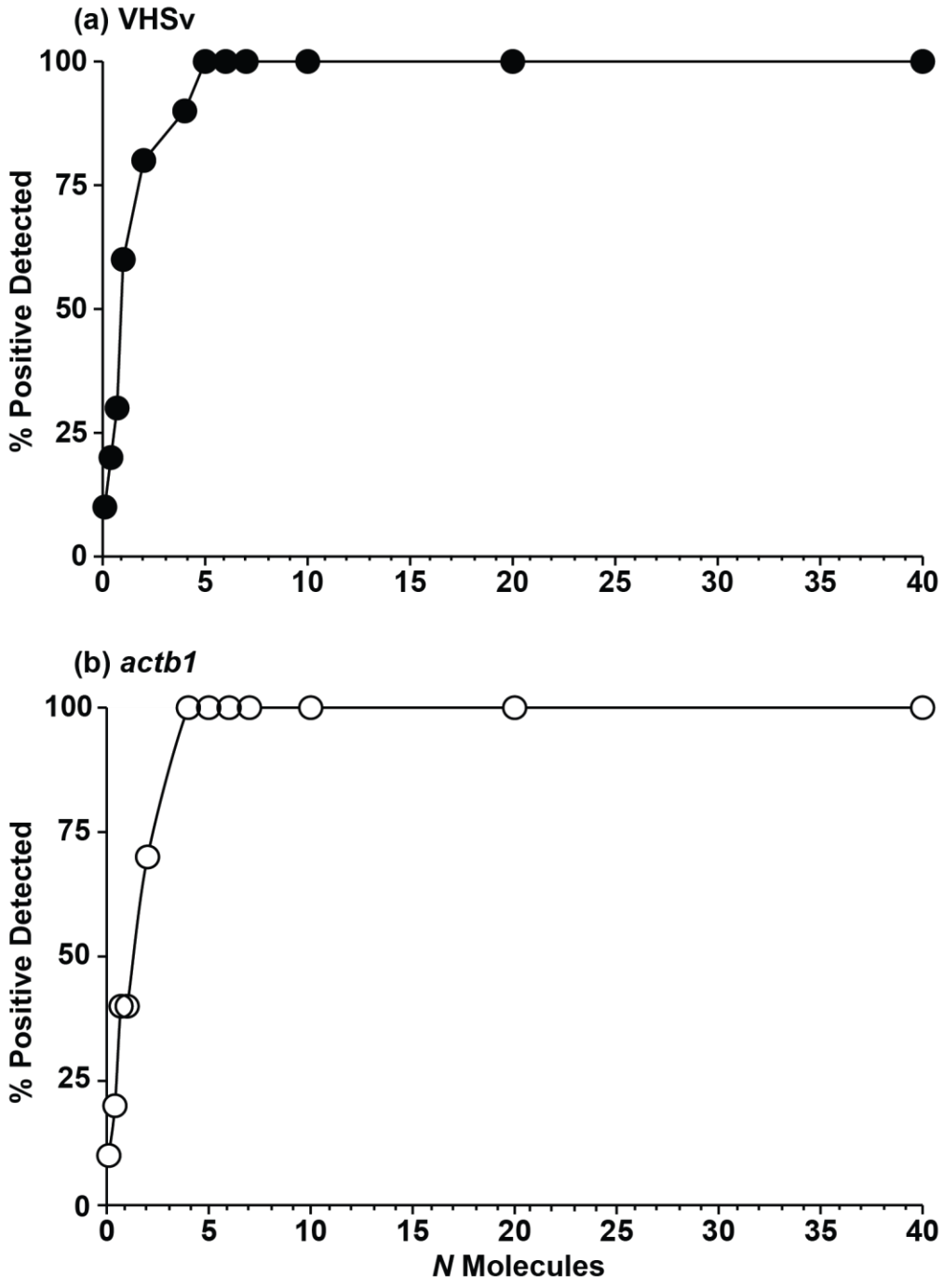

(a) VHSv
% Positive Detected
100
75
50
25
0
0
5
10
15
20
25
30
35
40
(b) actb1
% Positive Detected
100
75
50
25
0
0
5
10
15
20
25
30
35
40
N Molecules

Fig. 5-3

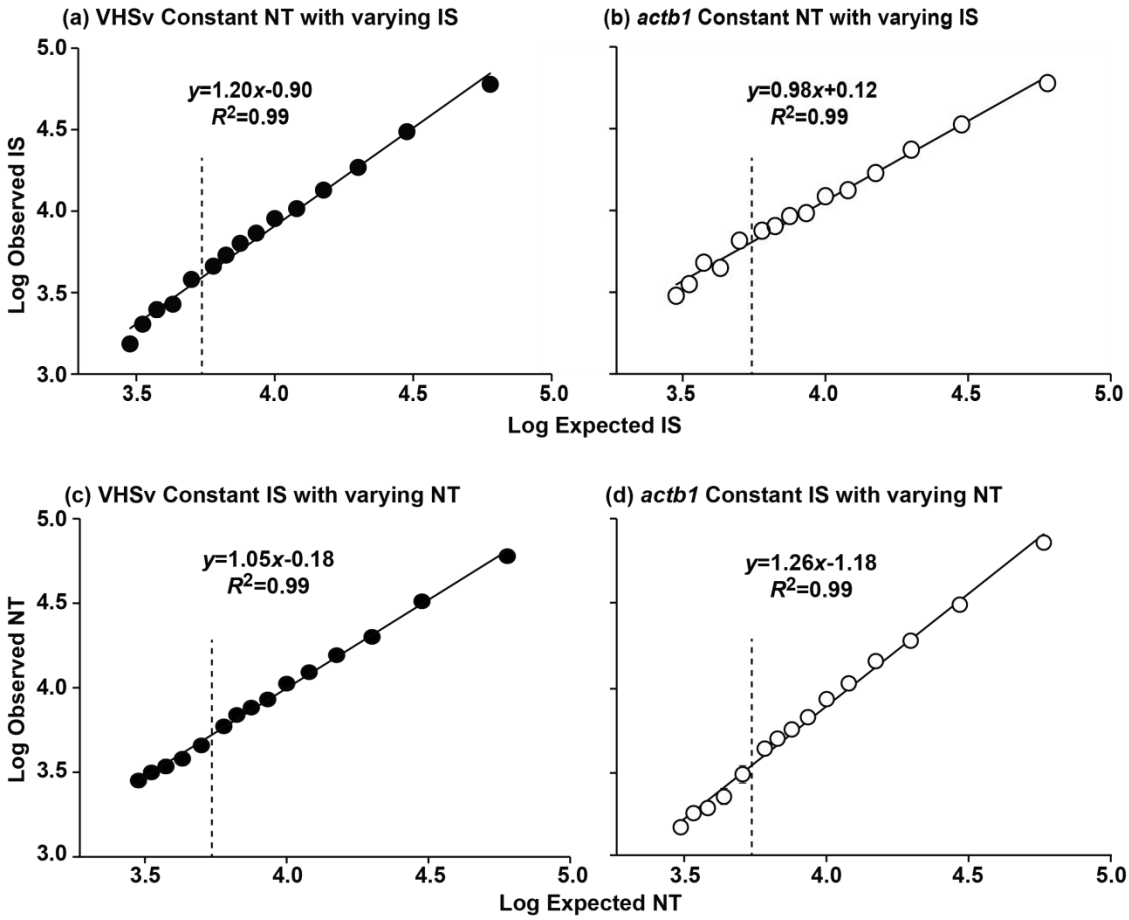

(a) VHSv Constant NT with varying IS
y=1.20x-0.90
R2=0.99
(b) actb1 Constant NT with varying IS
y=0.98x+0.12
R2=0.99
Log Observed IS
Log Expected IS
(c) VHSv Constant IS with varying NT
y=1.05x-0.18
R2=0.99
(d) actb1 Constant IS with varying NT
y=1.26x-1.18
R2=0.99
Log Observed NT
Log Expected NT
5.0
4.5
4.0
3.5
3.0

Fig. 5-4

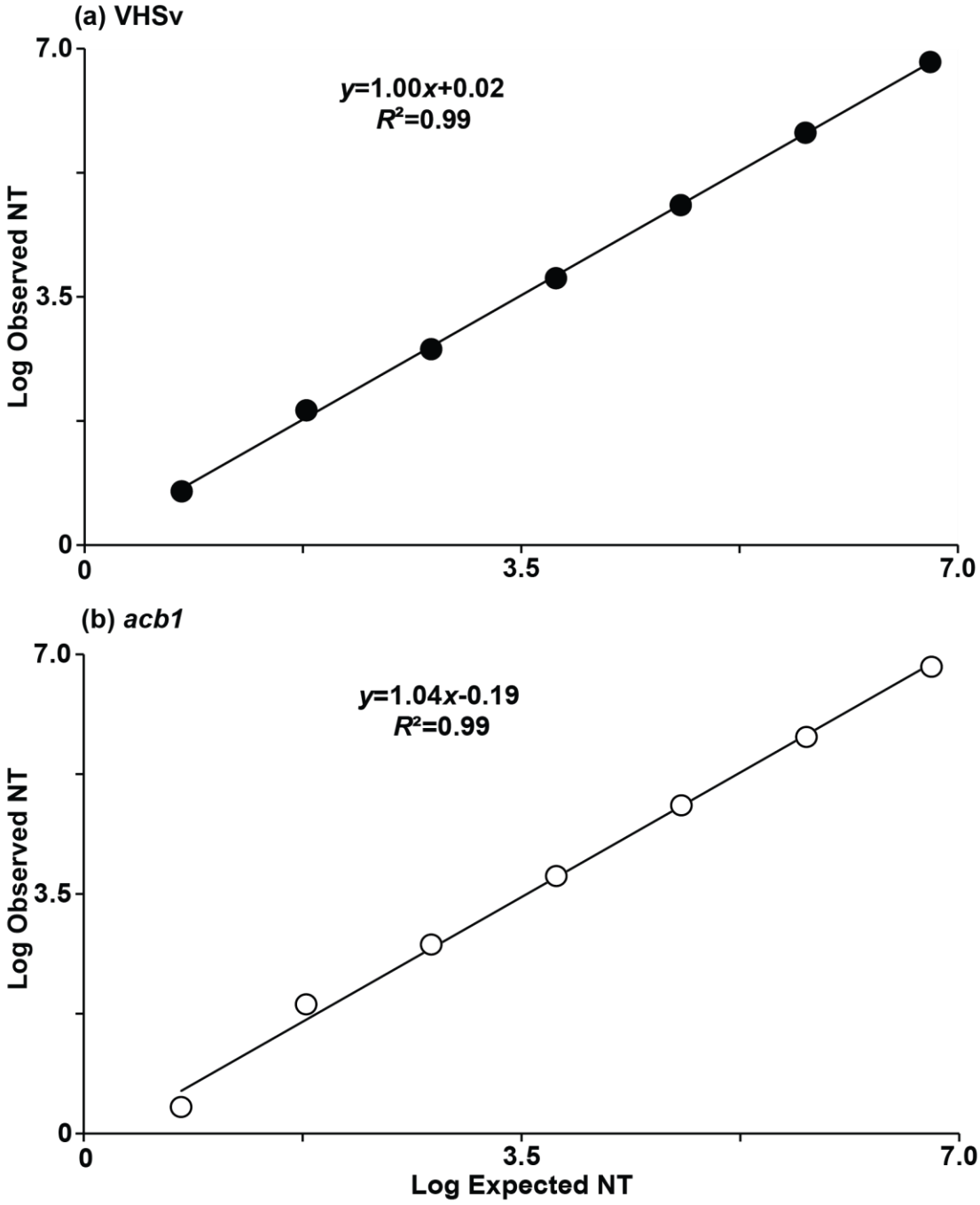

(a) VHSv
y=1.00x+0.02
R²=0.99
Log Observed NT
7.0
3.5
0
0
3.5
7.0
(b) acb1
y=1.04x-0.19
R²=0.99
Log Observed NT
7.0
3.5
0
0
3.5
7.0
Log Expected NT

Fig. 5-5

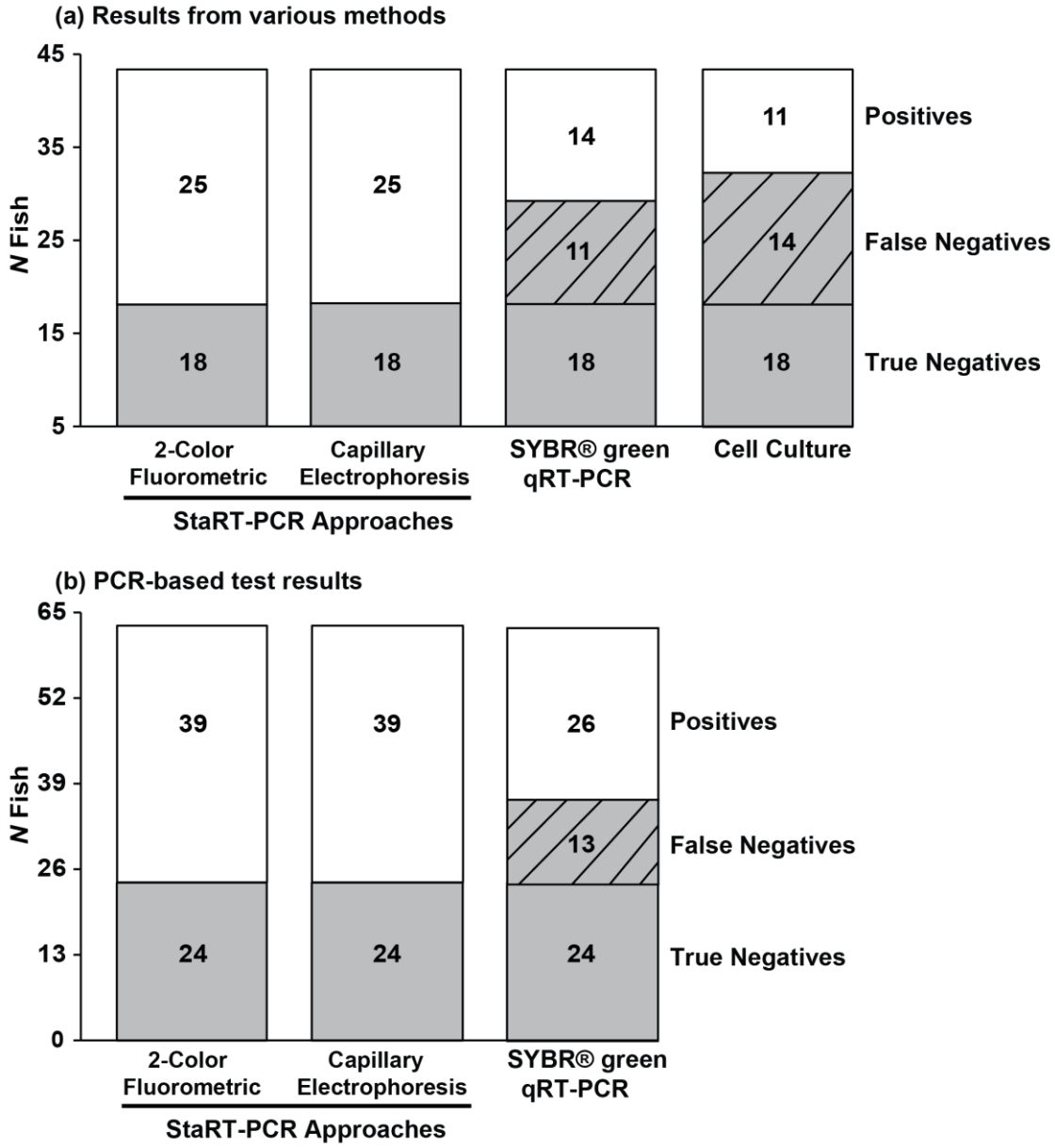
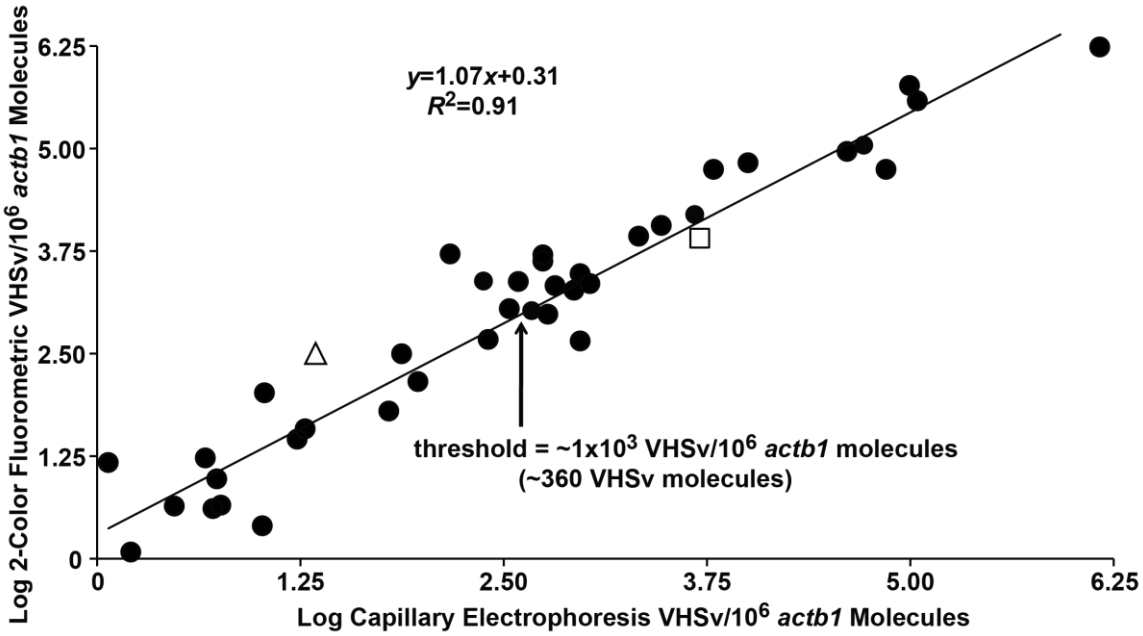

(a) Results from various methods
N Fish
45
35
25
15
5
25
25
14
11
11
14
18
18
18
18
Positives
False Negatives
True Negatives
2-Color Fluorometric
Capillary Electrophoresis
SYBR® green qRT-PCR
Cell Culture
StaRT-PCR Approaches
(b) PCR-based test results
N Fish
65
52
39
26
0
39
39
26
13
24
24
24
Positives
False Negatives
True Negatives
2-Color Fluorometric
Capillary Electrophoresis
SYBR® green qRT-PCR
StaRT-PCR Approaches

Fig. 5-6


y=1.07x+0.31
R2=0.91
threshold = ~1x10^3 VHSv/10^6 actb1 molecules
(~360 VHSv molecules)
Log 2-Color Fluorometric VHSv/10^6 actb1 Molecules
Log Capillary Electrophoresis VHSv/10^6 actb1 Molecules
0
1.25
2.50
3.75
5.00
6.25

Chapter 6

Discussion

6.1 General conclusions

This collaborative Ph.D. dissertation research investigates the evolutionary and biogeographic patterns of the Viral Hemorrhagic Septicemia virus (VHSv) and develops a new rapid and accurate assay to diagnose and quantify infection. Chapter 2 (published as Pierce and Stepien, 2012) evaluates the phylogenetic and biogeographic relationships among the four strains (I-IV) and substrains of VHSv in relation to a quasispecies diversification model. Genetic diversity and divergence times are compared among the glycoprotein (*G*), nucleoprotein (*N*), and non-virion (*Nv*) genes. Chapter 3 examines evolutionary rates, mutation types, and biogeographic distribution of VHSv substrain IVb in the Laurentian Great Lakes of North America in relation to possible gene function. Chapter 4 (published as Pierce et al., 2013) describes a new Standardized Reverse Transcriptase Polymerase Chain Reaction (StaRT-PCR) assay to detect and quantify known VHSv variants, comparing its performance and accuracy to conventional real-time quantitative (qRT) PCR-based tests (e.g., SYBR® green qRT-PCR) and cell culture (the "gold standard" for VHSv detection). Chapter 5 converts the StaRT-PCR assay from

Chapter 4 (published as Pierce and Willey et al., 2013) to a more user-friendly platform on common real-time PCR equipment. Qualitative and quantitative results from both platforms are compared. Overall, this dissertation study describes rapid diversification and spread of VHSv as a quasispecies, and develops a fast and accurate assay to detect its sequence variants.

6.1.1 VHSv evolution: Phylogeography and gene diversification

One of the primary conclusions arising from Chapters 2 and 3 is that VHSv is evolving rapidly, following a quasispecies model. Mutational diversification likely enhances the virus' ability to enter host cells and evade host immune responses, underlying its differentiation as four genetically and geographically distinct strains in the Northern Hemisphere (I-IV; Snow et al., 1999).

VHSv appears to have originated from a marine ancestor in the North Atlantic Ocean, diverging ~700 ya into two clades: strain IV in North America, and strains I-III in Europe. Strain II comprises the basal group of the latter clade, diverging ~570 ya in the Baltic Sea to infect marine fishes. I and III are sister strains, diverging ~400 ya, with the former mostly in European freshwaters and the latter in North Sea marine/estuarine waters. Strain IV diverged ~150 ya into three reciprocally monophyletic and geographically isolated substrains: IVa infects salmonids in the Northeastern Pacific, with some cases in Japan (the Japanese flounder *Paralichthys olivaceus*), IVb occurs in the freshwater Great Lakes, and IVc is found in estuarine fishes near New Brunswick, Canada (Chapter 2).

Broad-scale patterns suggest that substrains IVa and IVc may share a common ancestry in the Northwest Atlantic Ocean. The ancestral strain IV may have been transported across the Arctic from the Northwest Atlantic Ocean to the Northeast Pacific Ocean by a shipping or organismal vector. Alternatively, it may have been transported from the Atlantic coast in attempts to establish Atlantic salmon (*Salmo salar*) populations in the Pacific Northwest (Washington State Dept. of Health, 2011). It later appears to have been transported from the Pacific Northwest to Japan. Substrain IVb may have traveled down the St. Lawrence Seaway from the Northwest Atlantic Ocean to infect the Great Lakes.

Since VHSv-IVb was first identified in 2003 from a muskellunge (*Esox masquinongy*) in Lake St. Clair, multiple outbreaks were reported (2005, 2006, 2007, 2008, 2009 and 2011), with the largest in 2006 spanning across much of the Great Lakes. Over the past decade, IVb diversified into at least 19 unique *G*-gene variants (Groocock et al., 2007; Garver et al., 2013; MEAP-VHSv, 2013; Chapter 3), with Lake Erie appearing to house the greatest genetic diversity (8 of the 19 *G*-gene haplotypes; 42%). This may be due to its warm water temperatures, high fish abundance, and proximity to the original IVb identification site in Lake St. Clair.

Sequence analysis of 12 variants (haplotypes a-l) with unique *G*-gene haplotypes reveals they also have unique *Nv*-gene haplotypes, and less variation in the *P*-gene (4 haplotypes) and the *M*-gene (6 haplotypes). The *Nv* evolves the fastest ($2.0x10^{-3}$ nucleotide substitutions per site/year – μ), followed by the *G* ($\mu=2.1x10^{-4}$), *P* ($\mu=1.5x10^{-4}$), and *M* ($\mu=1.2x10^{-4}$). *Nv*'s rapid rate may be due to its role in suppressing apoptosis at early viral infection (Ammayappan and Vakharia, 2011), which may allow the virus to

persist longer in the host and evolve new mutations. This suggests that there are many more *Nv* variants that have yet to be discerned.

The *G*-gene evolves at ~1/10 the rate of *Nv*, which is similar to the overall rate published for VHSv as a whole ($2.6x10^{-4}$; Chapter 2), as well as IHNv ($2.0x10^{-4}$; Kurath et al., 2003) and rabies ($4.1x10^{-4}$; Holmes et al., 2002). It encodes the glycoprotein that forms spike-like projections on the viral particle; these attach to host cells and induce endocytosis (Bearzotti et al., 1999). Mutations in this gene may aid adherence and entry into the cell to thwart the host's ability to recognize the virus. The *P* and *M* genes evolve more slowly (14 and 17x less than *Nv*) and have fewer haplotypes. The functions of *P* and *M* thus appear more conserved; *P* plays a role in viral replication and inhibition of the host's innate immune response (Pore, 2012), and *M* blocks protein synthesis of infected host cells (Pore, 2012) and functions in viral budding (Biacchesi et al., 2002).

Among the 12 *G*-gene haplotypes, a (the original MI03GL isolate) and b (vcG002 from the 2005 Bay of Quinte, Lake Ontario outbreak) appear to be the oldest, being the most abundant (61 and 36) and widespread. Haplotype a spans all five Great Lakes and b is found in Lakes Michigan, Huron, Erie, and Ontario. A and b appear central in the *G*-gene haplotype network, with many other haplotypes diverging from them in a "star-like" (i.e., quasipecies) pattern. Haplotype b may have differentiated from a, after the first discovery of VHS-IVb in the Great Lakes. Alternatively, a and b appear relatively distant in the *Nv*-gene and may have had separate origins long prior to the outbreak.

Haplotypes k and l descend from a, which were collected from the same location (Sandusky OH, central Lake Erie) in 2012. K and l appear closely related to a sequences of all genes (in *P*, l retains the ancestral "a" sequence). It is interesting to note that the

fish with haplotypes k and l lacked clinical signs of VHSv infection (external hemorrhages), yet tested positive. Similar findings of fish samples that tested positive for VHSv, but lacked symptoms, were reported from monitoring (Kane-Sutton et al., 2010; Frattini et al., 2011) and other sequencing studies of the *G*-gene (Thompson et al., 2011; Cornwell et al., 2012; Faisal et al., 2012). These results suggest that although IVb has been evolving in the Great Lakes, it has become less virulent over time (i.e., massive fish die-offs have declined).

These findings imply that it may be important for future investigations to sequence a wider variety of isolates for the *Nv*-gene for fish samples that appear healthy and free of disease. Those results may further understanding of VHSv adaptive success and persistence.

6.1.2 Improved diagnostics for more sensitive and accurate VHSv detection

Cell culture is the sole USDA approved diagnostic to detect VHSv (OIE, 2013), which can take up to a month (Garver et al., 2011) and has a false negative range of 43-95% (Chico et al., 2006; López-Vázquez et al., 2006; Jonstrup et al., 2013). Most studies of viral surveillance in the Great Lakes have relied on cell culture (Kane-Sutton et al., 2010; Frattini et al., 2011; Cornwell et al., 2012), which may lead to misdiagnosis. Researchers have found that most fish that test positive for VHSv now lack clinical signs of the disease (i.e., external hemorrhages) (Thompson et al., 2011; Cornwell et al., 2012; Faisal et al., 2012), indicating that they may harbor levels of virus too low for cell culture to detect.

To circumvent this problem, we develop a unique StaRT-PCR assay to rapidly detect VHSv within a few hours, which includes synthetic internal standards (IS) to guard against false negative results (Chapter 4). Our assay originally was based on Agilent capillary electrophoresis, in which positive VHSv results are visualized as two peaks on the electropherogram (the positive control IS peak and the VHSv Native Template (NT) peak). However, since Agilent technology is uncommon in fish diagnostic laboratories, the assay was converted to use real-time PCR equipment (Chapter 5), which employs fluorescently labeled probes (true positives are visualized as two colors on the real-time PCR plot). Both of these assays are VHSv-specific, detect all four VHSv strains and substrains, and are free of false negatives. Other real-time qRT-PCR assays, which lacked IS, reported up to 92% false negatives (Chico et al., 2006; López-Vázquez et al., 2006; Liu et al., 2008; Matejusova et al., 2008; Cutrín et al., 2009; Hope et al., 2010; Garver et al., 2011; Jonstrup et al., 2013).

Our StaRT-PCR assays show superior detection, diagnosing infection down to a single VHSv molecule, with 100% accuracy at 5 molecules. In comparison, tests by Hope et al. (2010) and Garver et al. (2011) needed ≥100 viral copies, and an assay by Liu et al. (2008) required ≥140 VHSv molecules.

The StaRT-PCR assays control for interfering substances at the PCR level. Experiments with increasing amounts of EDTA (an inhibitor) in the PCR reaction show accurate detection and quantification of VHSv-IVb across a wide range of possible inhibition. Other qRT-PCR approaches likely would have resulted in false negatives.

6.2 Future research and recommendations

This dissertation examines the evolution and biogeography of an emerging viral quasispecies and describes a novel assay for its rapid and accurate detection. Outputs will enhance ability to detect the virus, predict and track its future spread, and serve as a baseline to interpret its evolutionary diversification pathways and mutation patterns. Several areas may be addressed in future studies, as follows:

6.2.1 VHSv-IVb evolution: Characterizing diversity

Chapter 3 analyzes sequences from the *G*, *Nv*, *P*, and *M* genes for 12 of the 19 with unique *G*-gene variants. Additional geographic sampling breadth is warranted. For example, the most recent IVb *G*-gene haplotypes were discovered by us from fish found in 2012 from Sandusky Bay, OH, indicating that this area (and others) likely harbor additional VHSv variants that have yet to be determined. Sequencing of the *N* and *L* genes is warranted to further understand mutational patterns. Analysis of the entire genome is recommended. This leads to questions, such as: (1) How many additional VHSv-IVb variants exist and where are they geographically distributed?, and (2) What are the respective functional roles of the VHSv genes and do variants differ in their respective fitness?

6.2.2 VHSv spread: Addressing the ballast water assumption

It has been hypothesized that VHSv may have been transported by ballast water through shipping to infect new waterbodies (Bain et al., 2010; Thompson et al., 2011). Testing ballast water at the top 10 Great Lakes exchange ports identified by the National Ballast Information Clearinghouse (NBIC, http://invasions.si.edu/nbic/) using the StaRT-PCR assay may address the question, does ballast water transport help spread VHSv?. Sequencing of positive samples will help to identify the pathways of variants, aiding understanding of the biogeographic diversification pathways of VHSv.

6.2.3 VHSv detection: Recommendations when using StaRT-PCR

In the StaRT-PCR assays, several standardized mixtures of internal standards (SMIS) were formulated, targeting low to high concentrations of VHSv molecules. Thus, there are 8 different SMIS to choose from when testing a fish sample. During this research, identifying the correct SMIS to use sometimes was problematic. For example, each fish may have one of three different diagnoses: (1) VHSv negative, (2) low numbers of VHSv molecules (1-1,000), and (3) high VHSv infection (>1,000 molecules). Since it is impossible to discern the level of infection by visual inspection, a random SMIS was selected for individual cases. If a fish was positive, but the resulting SMIS didn't amplify the NT and the IS concentrations within an order of magnitude of each other, the reaction needed to be re-run. Occasionally, the correct SMIS was identified in the first or second try, however, several attempts often were necessary, which leads to increased technician

time and reagent cost.

To improve the efficiency in selecting the correct SMIS, two different approaches may be evaluated: (1) in the pre-amplification reaction, a range of different SMIS mixtures could be simultaneously used to facilitate diagnosis across the realm of infection, or (2) a trial PCR run may evaluate the ratio of NT:IS *actb1* concentration prior to testing for VHSv. However, these two suggestions may consume larger amounts of test cDNA and PCR reagents. Thus, future researchers should evaluate a more efficient means to identify the correct SMIS.

References

Alonso M, Kim CH, Johnson MC, Pressley M, Leong J (2004) The *Nv* gene of Snakehead rhabdovirus (SHRv) is not required for pathogenesis, and a heterologous glycoprotein can be incorporated into the SHRv envelope. J Virol 78: 5875-5882

Ammayappan A, Kurath G, Thompson TM, Vakharia VN (2010) A reverse genetics system for the Great Lakes strain of Viral Hemorrhagic Septicemia virus: the *Nv* gene is required for pathogenicity. Mar Biotechnol 13: 672-683

Ammayappan A, Vakharia VN (2009) Molecular characterization of the Great Lakes Viral Hemorrhagic Septicemia virus (VHSv) isolate from USA. Virol J 6: 171 doi:10.1186/1743-422X-6-171

Ammayappan A, Vakharia VN (2011) Nonvirion protein of *Novirhabdovirus* suppresses apoptosis at the early state of virus infection. J Virol 85(16): 8393-8402

Aquatic Invasive Species Action Plan (2011) Viral Hemorrhagic Septicemia virus. Fish Production Services. Pennsylvania Fish and Boat Commission. Available from: http://fishandboat.com/ais/ais-action1-vhs.pdf

Bain MB, Cornwell ER, Hope KM, Eckerlin GE, Casey RN, Groocock GH, Getchell RG, Bowser PR, Winton JR, Batts WN, Cangelosi A, Casey JW (2010) Distribution of an invasive aquatic pathogen (Viral Hemorrhagic Septicemia virus) in the Great

Lakes and its relationship to shipping. PlOS One 5. doi:10.1371/journal.pone.0010156

Barbosa TFS, de Almeida Medeiros DB, da Rosa ES, Casseb LMN, Medeiros R, de Souza Pereira A, Vallinoto ACR, Vallinoto M, Begot AL, da Silva Lima RJ, da Costa Vasconcelos PF, Nunes MRT (2008) Molecular epidemiology of rabies virus isolated from different sources during a bat-transmitted human outbreak occurring in Augusto Correa municipality, Brazilian Amazon. Virol 2: 228-236

Basurco B, Benmansour A (1995) Distant strains of the fish rhabdovirus VHSv maintain a sixth functional cistron which codes for a nonstructural protein of unknown function. Virol 212: 741-745

Bearzotti M, Delmas B, Lamoureux A, Loustau AM, Chilmonczyk S, Bremont M (1999) Fish rhabdovirus cell entry is mediated by fibronectin. J Virol 73: 7703-7709

Belshaw R, Gardner A, Ranbaut A, Pybus OG (2008) Pacing a small cage: mutation and RNA viruses. Trends Ecol Evol 23: 188-193

Benmansour A, Basurco B, Monnier AF, Vende P, Winton JR, de Kinkelin P (1997) Sequence variation of the glycoprotein gene identifies three distinct lineages within field isolates of Viral Haemorrhagic Septicaemia virus, a fish rhabdovirus. J Gen Virol 78: 2837-2846

Benmansour A, Paubert G, Bernard J, De Kinkelin P (1994) The polymerase associated protein (*M*1) and the matrix protein (*M*2) from a virulent and an avirulent strain of Viral Hemorrhagic Septicemia virus (VHSv), a fish rhabdovirus. Virol 198(2): 602-612

Bernard J, Bremont M, Winton J (1992) Nucleocapsid gene sequence of a North American isolate of Viral Haemorrhagic Septicaemia virus, a fish rhabdovirus. J Gen Virol 73: 1011-1014

Bernard J, Lecocq-Xhonneux F, Rossius M, Thiry ME, de Kinkelin P (1990) Cloning and sequencing the messenger RNA of the *N* gene of Viral Haemorrhagic Septicaemia virus. J Gen Virol 71: 1669-1674

Betts AM, Stone DM (2000) Nucleotide sequence analysis of the entire coding regions of virulent and avirulent strains of Viral Haemorrhagic Septicaemia virus. Virus Genes 20: 259-262

Biacchesi S, Bearzotti M, Bouguon E, Bremont M (2002) Heterologous exchanges of the glycoprotein and the matrix protein in a *Novirhabdovirus*. J Virol 76(6): 2881-2889

Biacchesi S, Lamoureux A, Merour E, Bernard J, Bremont M (2010) Limited interference at the early state of infection between two recombinant Novirhabdoviruses: Viral Hemorrhagic Septicemia virus and Infectious Hematopoietic Necrosis virus. J Virol 84(19): 10038-10050

Biacchesi S, Thoulouze M, Bearzotti M, Yu Y, Bremont M (2000) Recovery of *Nv* knockout Infectious Hematopoietic Necrosis virus expressing foreign genes. J Virol 74(23): 11247-11253

Bowser P (2009) Fish Diseases: Viral hemorrhagic septicemia (VHS). Northeastern Regional Aquaculture Center. http://www.ca.uky.edu/wkrec/NRAC-VHS.pdf

Bruchhof B, Marquardt O, Enzmann PJ (1995) Differential diagnosis of fish pathogenic rhabdovirus by reverse transcriptase-dependent polymerase chain reaction. J Virol Methods 55: 111-119

Brunson R, True K, Yancey J (1989) VHS virus isolated at Makah National Fish Hatchery. American Fisheries Society, Fish Health Section Newsletter 17: 3-4

Burbrink FT, Pyron RA (2008) The taming of the skew: estimating proper confidence intervals for divergence dates. Syst Biol 57: 317-328

Bustin SA, Nolan T (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J Biomol Tech 15: 155-166

Byon JY, Takano T, Hirono I, Aoki T (2006) Genome analysis of Viral Hemorrhagic Septicemia virus isolated from Japanese flounder *Paralichthys olivaceus* in Japan. Fisheries Sci 72: 906-908

Campbell S, Collet B, Einer-Jensen K, Secombes CJ, Snow M (2009) Identifying potential virulence determinants in Viral Haemorrhagic Septicaemia virus (VHSv) for rainbow trout. Dis Aquat Org 86: 205-212

Canales RD, Luo Y, Willey JC, Austermiller B, Barbacioru CC, Boysen C, Hunkapiller K, Jensen RV, Knight CR, Lee KY, Ma Y, Maqsodi B, Papallo A, Peters EH, Poulter K, Ruppel PL, Samaha RR, Shi L, Yang W, Zhang L, Goodsaid FM (2006) Evaluation of DNA microarray results with quantitative gene expression platforms. Nat Biotech 24: 1115-1122

Celi FS, Zenilman ME, Shuldiner AR (1993) A rapid and versatile method to synthesize internal standards for competitive PCR. Nucleic Acids Res 21: 1047

Chai Z, Ma W, Fu F, Lang Y, Wang W, Tong G, Liu Q, Cai X, Li X (2013) A SYBR

Green-based real-time RT-PCR assay for simple and rapid detection and differentiation of highly pathogenic and classical type 2 porcine reproductive and respiratory syndrome virus circulating in China. Arch Virol 158: 407-415

Chico V, Gomez N, Estepa A, Perez L (2006) Rapid detection and quantitation of Viral Hemorrhagic Septicemia virus in experimentally challenged rainbow trout by real-time PCR. J Virol Methods 132: 154-159

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9: 1657-1659

Cohen J (1992) A primer power. Quant Methods Psych 1: 155-159

Cornwell ER, Eckerlin GE, Thompson TM, Batts WN, Getchell RG, Groocock GH, Kurath G, Winton JR, Casey RN, Casey JW, Bain MB, Bowser PR (2012) Predictive factors and viral genetic diversity for Viral Hemorrhagic Septicemia virus infection in Lake Ontario and the St. Lawrence River. J Great Lakes Res 38: 278-288

Coutlée F, Viscidi RP, Saint-Antoine P, Kessous A, Yolken RH (1991) The polymerase chain reaction: a new tool for the understanding and diagnosis of HIv-1 infection at the molecular level. Mol Cell Probe 5: 241-259

Crawford EL, Warner KA, Khuder SA, Zahorchak RJ, Willey JC (2002) Multiplex standardized RT-PCR for expression analysis of many genes in small samples. Biochem Biophys Res Comm 293: 509-516

Cuevas JM, Duffy S, Sanjuán R (2009) Point mutation rates of bacteriophage ΦX174. Genetics 747-749

Cunningham CW (1997) Can three incongruence tests predict when data should be combined? Mol Biol Evol 14: 733-740

Cutrín JM, Olveira JG, Bandín I, Dopazo CP (2009) Validation of real-time RT-PCR applied to cell culture for diagnosis of any known genotype of Viral Haemorrhagic Septicaemia virus. J Virol Methods 162: 155-162

Dale OB, Ørpetveit I, Lyngstad TM, Kahns S, Skall HF, Olesen NJ, Dannevig BH (2009) Outbreak of Viral Haemorrhagic Septicaemia (VHS) in seawater-farmed rainbow trout in Norway caused by VHS virus genotype III. Dis Aqua Organ 85: 93-103

Dietzgen RG (2012) Morphology, genome organization, transcription and replication of rhabdoviruses. In: Dietzgen RG, Kuzmin IV (eds.), Rhabdoviruses, Molecular Taxonomy, Evolution, Genomics, Ecology, Host-Vector Interactions, Cytopathology, and Control, Caister Academic Press, Norfolk, United Kingdom, pp. 5-11

Dixon PF (1999) VHSv came from the marine environment: clues from the literature, or just red herrings? Bull Eur Assoc Fish Pathol 19: 60-65

Domingo E (2006) Virus evolution, In: Knipe DM, Howley PM (eds.), Fields Virology, 5th edition. Lippincott Williams and Wilkins, Baltimore, Maryland, pp. 389-406

Domingo E, Biebricher CK, Eigen M, Holland JJ (2002) Quasispecies and RNA virus evolution, In: Principles and Consequences. Landes Bioscience New York, USA

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLOS Bio 4: e88. doi:10.1371/journal.pbio.0040088

Drummond AJ, Nocholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally

spaced sequence data. Genetics 161: 1307-1320

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUTi and the BEAST 1.7. Mol Biol Evol 29(8)1969-1973. Available from: http://beast.bio.ed. ac.uk/Main_Page. [Assessed 30 July 2013]

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

Einer-Jensen K, Ahrens P, Forsberg R, Lorenzen N (2004) Evolution of the fish rhabdovirus Viral Haemorrhagic Septicaemia virus. J Gen Virol 85: 1167-1179

Einer-Jensen K, Ahrens P, Lorenzen N (2005) Parallel phylogenetic analyses using the *N*, *G* or *Nv* gene from a fixed group of VHSv isolates reveal the same overall genetic typing. Dis Aqua Organ 67: 39-45

Elena SF, Sanjuán R (2005) Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J Virol 79: 11555-11558

Ellis JS, Zambon MC (2002) Molecular diagnosis of influenza. Rev Med Virol 12: 375-389

Elsayed E, Faisal M, Thomas M, Whelan G, Batts W, Winton J (2006) Isolation of Viral Haemorrhagic Septicaemia virus from muskellunge, *Esox masquinongy* (Mitchill), in Lake St Clair, Michigan, USA reveals a new sublineage of the North American genotype. J Fish Dis 29: 611-619

Enzmann PJ, Castric J, Bovo G, Thiéry R, Fichtner D, Schutze H, Wahli T (2010) Evolution of Infectious Hematopoietic Necrosis virus (IHNv), a fish rhabdovirus, in Europe over 20 years: implications for control. Dis Aqua Organ 89: 9-15

Erdfelder E, Faul F, Buchner A (1996) GPOWER: a general power analysis program. Behav Res Methods Instrum Comp 28: 1-11

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10: 564-567 Available from: http://cmpg.unibe.ch/software/arlequin35/. [Accessed 30 July 2013]

Faisal M, Schulz CA (2009) Detection of Viral Hemorrhagic Septicemia virus (VHSv) from the leech *Myzobdella lugubris* Leidy, 1851. Parasite Vector 2: 45. doi:10.1186/1756-3305-2-45

Faisal M, Shavalier M, Kim RK, Millard EV, Gunn MR, Winters AD, Schulz CA, Eissa A, Thomas MV, Wolgamood M, Whelan GE, Winton J (2012) Spread of emerging Viral Hemorrhagic Septicemia virus strain, genotype IVb, in Michigan, USA. Viruses 4: 734-760

Faisal M, Winters AD (2011) Detection of viral hemorrhagic septicemia virus (VHSv) from *Diporeia* spp. (Pontoporeiidae, Amphipoda) in the Laurentian Great Lakes, USA. Parasite Vector 4: 2. doi:10.1186/1756-3305-4-2

Farris J S, Kallersjo M, Kluge AG, Bult C (1994) Testing significance of congruence. Cladistics 10: 315-320

Farris JS, Källersjö M, Kluge AG, Bult C (1995) Constructing a significance test for incongruence. Syst Biol 44: 570-572

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791

Fontoura BM, Faria PA, Nussenzveig DR (2005) Viral interactions with the nuclear transport machinery: discovering and disrupting pathways. Life 57(2): 65-72

Frattini SA, Groocock GH, Getchell RG, Wooster GA, Casey RN, Casey JW, Bowser PR (2011) A 2006 survey of Viral Hemorrhagic Septicemia (VHSv) virus type IVb in New York state waters. J Great Lakes Res 37: 194-198

Gadd T, Jakava-Viljanen M, Einer-Jensen K, Ariel E, Koski P, Sihvonen L (2010) Viral Haemorrhagic Septicaemia virus (VHSv) genotype II isolated from European river lamprey *Lampetra fluviatilis* in Finland during surveillance from 1999 to 2008. Dis Aquat Org 17: 189-198

Gadd T, Jakava-Viljanen M, Tapiovaara H, Koski P, Sihvonen L (2011) Epidemiological aspects of Viral Haemorrhagic Septicaemia virus genotype II isolated from Baltic herring, *Clupea harengus membras L.* J Fish Dis 34: 517-529

Gagné N, MacKinnon AM, Boston L, Souter B, Cook-Versloot M, Griffiths S, Olivier G (2007) Isolation of Viral Haemorrhagic Septicaemia virus from mummichog, stickleback, striped bass and brown trout in eastern Canada. J Fish Dis 30: 213-223

Galinier R, van Beurden S, Amilhat E, Castric J, Schoehn G, Verneau O, Fazio G, Allienne JF, Engelsma M, Sasal P, Faliex E (2012) Complete genomic sequence and taxonomic position of the eel virus European X (EVEX), a rhabdovirus of European eel. Virus Res 166: 1-12

Garver KA, Hawley LM, McClure CA, Schroeder T, Aldous S, Doig F, Snow M, Edes S, Baynes C, Richard J (2011) Development and validation of a reverse transcription quantitative PCR for universal detection of Viral Hemorrhagic Septicemia virus. Dis Aquat Organ 95: 97-112

Garver KA, Traxler GS, Hawley LM, Richard J, Ross JP, Lovy J (2013) Molecular epidemiology of Viral Haemorrhagic Septicaemia virus (VHSV) in British Columbia, Canada, reveals transmission from wild to farmed fish. Dis of Aquat Organ 104: 93-104.

Gelderblom HC, Menting S, Beld MG (2006) Clinical performance of the new Roche COBAS (R) TaqMan HCv test and high pure system for extraction, detection and quantification of HCv RNA in plasma and serum. Antivir Ther 11: 95-103

Gilliland G, Perrin S, Blanchard K, Bunn HF (1990) Analysis of cytokine mRNA and DNA: detection and quantitation by competitive polymerase chain reaction. Proc Nat Acad Sci 87: 2725-2729

Gillou JP, Merle G, Henault S, Hattenberger AM (1999) Detection of Viral Hemorrhagic Septicemia virus (VHSv) in rainbow trout (*Oncorhynchus mykiss*) by reverse transcription followed by polymerase chain reaction. Diagnostic validation. Vet Res 30: 49-60

Goodwin AE, Merry GE (2011) Replication and persistence of VHSv IVb in freshwater turtles. Dis Aquat Org 94: 173-177

Grady W (2007) The freshwater seas, In: The Great Lakes: The Natural History of a Changing Region. Greystone Books, Vancouver, British Columbia, Canada, pp. 5-38

Great Lakes Fish Health Committee (2011) Annual report to the GLFHC. Wisconsin Department of Natural Resources. Available from: http://www.glfc.org/ boardcomm/fhealth/ 2011annualreport.pdf. [Accessed 20 July 2013]

Groocock GH, Getchell RG, Wooster GA, Britt KL, Batts WN, Winton JR, Casey RN, Casey JW, Bowser PR (2007) Detection of Viral Hemorrhagic Septicemia in round gobies in New York State (USA) waters of Lake Ontario and the St. Lawrence River. Dis Aquat Org 76: 187-192

Guindon S, Gascuel O (2003) PhyML3.0: a simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52(5): 696-704. Available from: http://atgc.lirmm.fr/phyml/. [Accessed 30 July 2013]

Hall N, Karras M, Raine JD, Carlton JM, Kooij TWA, Berriman M, Florens L, Janssen CS, Pain A, Christophides GK, James K, Rutherford K, Harris B, Harris D, Churcher C, Quail MA, Ormond D, Doggett J, Trueman HE, Mendoza J, Bidwell SL, Rajandream M, Carucci DJ, Yates III JR, Kafatos FC, Janse CJ, Barrell B, Turner MR, Waters AP, Sinden RE (2005) A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. Science 307: 82-86

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98. Available from: http://www.mbio.ncsu.edu/ bio edit/page2.html. [Accessed 30 July 2013]

Hawley LM, Garver KA (2008) Stability of Viral Hemorrhagic Septicemia virus (VHSv) in freshwater and seawater at various temperatures. Dis Aquat Org 82: 171-178

Hedrick RP, Batts WN, Yun S, Traxler GS, Kaufman J, Winton JR (2003) Host and geographic range extensions of the North American strain of Viral Hemorrhagic Septicemia virus. Dis Aquat Org 55: 211-220

Hicks AL, Duffy S (2011) Genus-specific substitution rate variability among Picornaviruses J Virol 85(15): 7942-7947

Holmes EC, Woelk CH, Kassis R, Bourhy H (2002) Genetic constraints and the adaptive evolution of rabies virus in nature. Virol 292: 247-257

Holmes EC (2009) The Evolution and Emergence of RNA Viruses. Oxford University Press, Oxford, New York.

Hope KM, Casey RN, Groocock GH, Getchell RG, Bowser PR, Casey JW (2010) Comparison of quantitative RT-PCR with cell culture to detect Viral Hemorrhagic Septicemia virus (VHSv) IVb infections in the Great Lakes. J Aquat Anim Health 22: 50-61

Hopper K (1989) The isolation of VHSv from chinook salmon at Glenwood Springs, Orcas Island, Washington. American Fisheries Society, Fish Health Section Newsletter 17: 1-2

Hudson RR, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. Genetics 116: 153-159

Huggett JF, Novak T, Garson JA, Green C, Morris-Jones SD, Miller RF, Zumla A (2008) Differential susceptibility of PCR reactions to inhibitors: an important and unrecognised phenomenon. BMC Res Notes 1 (70). doi:10.1186/1756-0500-1-70

Hughes GJ, Orciari LA, Rupprecht CE (2005) Evolutionary timescale of rabies virus adaptation to North American bats inferred from the substitution rate of the nucleoprotein gene. J Gen Virol 86: 1467-1474

International Committee on Taxonomy of Viruses (2013) Virus taxonomy. Available from: http://www.ictvonline.org/virusTaxonomy.asp. [Accessed 30 July 2013]

International Organization for Standardization (2005) Microbiology of food and animal feeding stuffs-Polymerase chain reaction (PCR) for the detection of food-borne pathogens – General requirements and definitions. BS EN ISO 22174

Ito T, Kurita J, Sano M, Skall HF, Lorenzen N, Einer-Jensen K, Olesen NJ (2012) Typing of Viral Hemorrhagic Septicemia virus by monoclonal antibodies. J Gen Virol 93: 2546-2557

Jaag HM, Nagy PD (2010) The combined effect of environmental and host factors on the emergence of viral RNA recombinants. PlOS Path 6(10): doi:10.1371/journal. ppat.1001156

Johnson MC, Simon BE, Kim CH, Leong JA (2000) Production of recombinant Snakehead rhabdovirus: the *Nv* protein is not required for viral replication. J Virol 74: 2343-2350

Jonstrup SP, Kahns S, Skall HF, Boutrup TS, Olesen NJ (2013) Development and validation of a novel Taqman-based real-time RT-PCR assay suitable for demonstrating freedom from Viral Haemorrhagic Septicaemia virus. J Fish Dis 36: 9-23

Jukes TH, Cantor CR (1969) Evolution of protein molecules, In: Munro, H.N. (ed.), Mammalian Protein Metabolism. Academic Press, New York, pp. 21-132

Kane-Sutton M, Kinter B, Dennis PM, Koonce JF (2010) Viral Hemorrhagic Septicemia virus in yellow perch, *Perca flavescens*, in Lake Erie. J Great Lakes Res 36: 37-43

Kim R, Faisal M (2011) Emergence and resurgence of the Viral Hemorrhagic Septicemia virus (*Novirhabdovirus*, *Rhabdoviridae*, *Mononegavirales*). J Advan Res 2: 9-23

Kim RK, Faisal M (2010) The Laurentian Great Lakes strain (MI03) of the Viral Haemorrhagic Septicemia virus is highly pathogenic for juvenile muskellunge, *Esox masquinongy* (Mitchell). J Fish Dis 33: 513-527

Kim RK, Faisal M (2012) Shedding of Viral Hemorrhagic Septicemia virus (genotype IVb) by experimentally infected muskellunge (*Esox masquinongy*). J Microbiol 50: 278-284

Kimura M (1981) Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Natl Acad Sci 78: 454-458

Kinnear M, Linde CC (2010) Capsid gene divergence in rabbit hemorrhagic disease virus. J Gen Virol 91: 174-181

Korber B (2000) HIv signature and sequence variation analysis, In: Rodrigo, A.G., Learn, G.L., (eds.), Computational Analysis of HIv Molecular Sequences. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 55-72

Kurath G (2012) Fish novirhabdoviruses. In: Dietzgen RG, Kuzmin IV (eds.), Rhabdoviruses, Molecular Taxonomy, Evolution, Genomics, Ecology, Host-Vector Interactions, Cytopathology, and Control. Caister Academic Press, Norfolk, United Kingdom, pp. 89-116

Kurath G, Garver KA, Troyer RM, Emmenegger EJ, Einer-Jensen K, Anderson ED (2003) Phylogeography of Infectious Haematopoietic Necrosis virus in North America. J Gen Virol 84: 803-814

Kurath G, Higman KH, Björklund, HV (1997) Distribution and variation of *Nv* genes in fish rhabdoviruses. J Gen Virol 78: 113-117

Kuzmin IV, Hughes GJ, Rupprecht CE (2006) Phylogenetic relationships of seven previously unclassified viruses within the family *Rhabdoviridae* using partial nucleoprotein gene sequences. J Gen Virol 87: 2323-2331

Lanave C, Preparata G, Saccone C, Serio G (1984) A new method for calculating evolutionary substitution rates. J Mol Evol 20: 86-93

LaPatra SE, Evilia C, Winston V (2008) Positively selected sites on the surface glycoprotein (*G*) of Infectious Hematopoietic Necrosis virus. J Gen Virol 89: 703-708

Lauring AS, Andino R (2010) Quasispecies theory and the behavior of RNA viruses. PLOS Path 6. doi:10.1371/journal.ppat.1001005

Leighton FA (2011) Wildlife pathogens and diseases in Canada. Canadian Councils of Resource Ministers, Ottawa, Ontario, Canada

Leland DS, Ginocchio CC (2007) Role of cell culture for virus detection in the age of technology. Clin Micro Rev 20(1): 49-78

Li WH (1997) Molecular Evolution. Sinauer Associates, Sunderland, Maryland

Librado P, Rozas J (2009) DNAsp v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformations 25: 1451-1452. Available from: http://www.ub.edu /dnasp/. [Accessed 30 July 2013]

Liu Z, Teng Y, Liu H, Jiang Y, Xie X, Li H, Lv J, Gao L, He J, Shi X, Tian F, Yang J, Xie C (2008) Simultaneous detection of three fish rhabdoviruses using multiplex real-time quantitative RT-PCR assay. J Virol Methods 149: 103-109

López-Vázquez C, Raynard RS, Bain N, Snow M, Bandín I, Dopazo CP (2006) Genotyping of marine Viral Haemorrhagic Septicaemia virus isolated from the Flemish Cap by nucleotide sequence analysis and restriction fragment length polymorphism patterns. Dis Aquat Org 73: 23-31

Lou M, Green TJ, Zhang X, Tsao J, Qiu S (2007) Conserved characteristics of the rhabdovirus nucleoprotein. Virus Res 129: 246-251

Lumsden JS, Morrison B, Yason C, Russell S, Young K, Yazdanpanah A, Huber P, Al-Hussinee L, Stone D, Way K (2007) Mortality event in freshwater drum *Aplodinotus grunniens* from Lake Ontario, Canada, associated with Viral Haemorrhagic Septicemia virus, type IV. Dis Aquat Organ 76: 99-111

Makarewicz JC, Bertram P (1991) Evidence for the restoration of the Lake Erie ecosystem: water quality, oxygen levels, and pelagic function appear to be improving. BioScience 41(4): 216-223

Masatani T, Ito N, Shimizu K, Ito Y, Nakagawa K, Sawaki Y, Koyama H, Sugiyama M (2010) Rabies virus nucleoprotein functions to evade activation of the RIG-I-mediated antiviral response. J Virol 84: 4002-12

Matejusova I, McKay P, McBeath AJ, Collet B, Snow M (2008) Development of a sensitive and controlled real-time RT-PCR assay for Viral Haemorrhagic Septicaemia virus (VHSv) in marine salmonid aquaculture. Dis Aquat Org 80: 137-144

May FJ, Davis CT, Tesh RB, Barrett ADT (2011) Phylogeography of West Nile virus: from the cradle of evolution in Africa to Eurasia, Australia, and the Americas. J Virol 85: 2964-2974

McCord B, Opel K, Funes M, Zoppis S, Meadows JL (2011) An investigation of the effect of DNA degradation and inhibition on PCR amplification of single source and mixed forensic samples. Project 2006-DN-BX-K006. Final Report to the U.S. Department of Justice. Available from: https://www.ncjrs.gov/pdffiles1/nij/grants/236692.pdf. [Accessed 30 July 2013]

Meyers TR, Short S, Lipson K (1999) Isolation of the North American strain of Viral Hemorrhagic Septicemis virus (VHSv) associated with epizootic mortality in two new host species of Alaskan marine fish. Dis Aqua Org 38: 81-86

Meyers TR, Winton JR (1995) Viral Hemorrhagic Septicemia virus in North America. Ann Rev Fish Dis 5: 3-24

Millard EV, Faisal M (2012) Development of neutralizing antibody responses in muskellunge, *Esox masquinongy* (Mitchell), experimentally exposed to Viral Haemorrhagic Septicaemia virus (genotype IVb). J Fish Dis 35: 11-18

Miller TA, Rapp J, Wastlhuber U, Hoffmann RW, Enzmann PJ (1998) Rapid and sensitive reverse transcriptase polymerase chain reaction based detection and differential diagnosis of fish pathogenic rhabdoviruses in organ samples and cultured cells. Dis Aquat Organ 34: 13-20

Molecular epidemiology of aquatic pathogens – Viral Hemorrhagic Septicemia virus (MEAP-VHSv) (2013) Available from: http://gis.nacse.org /vhsv/#. [Accessed 30 July 2013]

Nichol ST, Rowe JE, Winton JR (1995) Molecular epizootiology and evolution of the glycoprotein and non-virion protein genes of infectious hematopoietic necrosis virus, a fish rhabdovirus. Virus Res 38: 159-173

Nishizawa T, Savaş H, Işıdan H, Üstündağ, C, Iwamoto H, Yoshimizu M (2006) Genotyping and pathogenicity of Viral Hemorrhagic Septicemia virus from free-living turbot (*Psetta maxima*) in a Turkish coastal area of the Black Sea. Appl Environ Microbiol 72: 2373-2378

Norusis M (2008) SPSS 16.0 Guide to Data Analysis. Prentice Hall Press, Upper Saddle River, New Jersey. Available from: http://www.spss.com. [Accessed 30 July 2013]

Ogino S, Kawakaki T, Brahmandam M, Cantor M, Kirkner GJ, Spiegelman D, Makrigiorgos GM, Weisenberger DJ, Laird PW, Loda M, Fuchs CS (2006) Precision and performance characteristics of bisulfite conversion and real-time PCR (MethyLight) for quantitative DNA methylation analysis. J Mol Diagn 8: 209-217

Ohio Division of Wildlife (2001) Lake Erie Strategic Plan. Available from: http://www.dnr.state.oh.us/Home/FishingSubhomePage/fisheriesmanagement placeholder/fishingfairportstratplan/tabid/6167/Default.aspx. [Accessed 30 July 2013]

Ojosnegros S, Beerenwinkel N (2010) Models of RNA virus evolution and their roles in vaccine design. Immun Res 6. doi:10.1186/1745-7580-6-S2-S5

OIE (Office of International des Epizooties), World Organization for Animal Health (2009) Viral Haemorrhagic Septicaemia, In: Vallet and Pastoret, P. (eds.), Manual

of Diagnostic Tests for Aquatic Animals. Office International des Epizooties, Paris, France. pp. 279-298

OIE, World Organization for Animal Health (2013) Viral Haemorrhagic Septicaemia. In: Olesen, N.J. and Skall, H.F. (eds.), Manual of Diagnostic Tests for Aquatic Animals. Office International des Epizooties, Paris, France, pp 347-396

Opel KL, Chung D, McCord BR (2010) A study of PCR inhibition mechanisms using real time PCR. J Forensic Sci 55: 25-33.

Oregon's Agricultural Progress (2011) Invading species. What lurks in eastern Oregon waterways? Extension and Experiment Station Communications, Oregon State University, Oregon. Available from: http://oregonprogress.oregonstate.edu/winter-2011/invading-species. [Accessed 30 July 2013]

Padhi A, Verghese B (2008) Detecting molecular adaptation at individual codons in the glycoprotein gene of the geographically diversified Infectious Hematopoietic Necrosis virus, a fish rhabdovirus. Virus Res 132: 229-236

Pal C, Maciá MD, Oliver A, Schachar I, Buckling A (2007) Coevolution with viruses drives the evolution of bacterial mutation rates. Nature 450: 1079-1081

Park S, Zhang Y, Lin S, Wang TH, Yang S (2011) Advances in microfluidic PCR for point-of-care infectious disease diagnostics. Biotech Adv 29: 830-839

Pasche JS, Mallik I, Anderson NR, Gudmestad NC (2013) Development and validation of a real-time PCR assay for the quantification of *Verticillium dahliae* in potato. Plant Dis 97: 608-618

Phelps NBD, Patnayak DP, Jiang Y, Goyal SM (2012) The use of a one-step real-time reverse transcription polymerase chain reaction (rRT-PCR) for the surveillance of

Viral Hemorrhagic Septicemia virus (VHSv) in Minnesota. J Aquat Anim Health 24: 238-243

Pierce LR, Stepien CA (2012) Evolution and biogeography of an emerging quasispecies: diversity patterns of the fish Viral Hemorrhagic Septicemia virus (VHSv). Mol Phylogenet Evol 63: 327-341

Pierce LR, Willey JC, Crawford EL, Palsule VV, Leaman DW, Faisal M, Kim RK, Shepherd BS, Stanoszek LM, Stepien CA (2013) A new StaRT-PCR approach to detect and quantify fish Viral Hemorrhagic Septicemia virus (VHSv): enhanced quality control with internal standards. J Virol Methods 189:129-142

Pierce LR, Willey JC, Palsule VV, Yeo J, Shepherd BS, Crawford EL, Stepien CA (2013) Accurate detection and quantification of the fish Viral Hemorrhagic Septicemia virus (VHSv) with a two-color fluorometric real-time PCR assay. PLOS One (in press)

Plemper RK (2011) Cell entry of enveloped viruses. Curr Opin Virol 1(2): 92-100

Pore AJ (2012) Studies on host-virus interaction for Viral Hemorrhagic Septicemia virus (VHSv). Master's thesis. Retrieved from OhioLink Dissertations and Theses. Available from: https://etd.ohiolink.edu/ap:10:0::NO:10:P10_ACCESSION_NUM:toledo 1336766667. [Accessed 30 July 2013]

Posada D (2008) jModelTest2: phylogenetic model averaging. Mol Biol Evol 25: 1253-1256. Available from: http://code.google.com/p/jmodeltest2/. [Accessed 30 July 2013]

Quer J, Huerta R, Novella IS, Tsimring L, Domingo E, Holland JJ (1996) Reproducible nonlinear population dynamics and critical points during replicative competitions

of RNA virus quasispecies. J Mol Biol 264: 465-471

R Development Core Team (2012) R: A language and environment for statistical computing, reference index version 2.13.1. R Foundation for Statistical Computing, Vienna, Austria. Available from: http://www.r-project.org/. [Accessed 30 July 2013]

Raja-Halli M, Vehmas TK, Rimaila-Pärnänen E, Sainmaa S, Skall HF, Olesen NJ, Tapiovaara H (2006) Viral Haemorrhagic Septicaemia (VHS) outbreaks in Finnish rainbow trout farms. Dis Aqua Organ 72: 201-211

Rao JR, Fleming CC, Moore JE (2006) Molecular Diagnostics Current Technology and Applications. Horizon Bioscience, Norfolk, Norwich, United Kingdom

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43: 223-225

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9: 552-569

Ronquist F, Huelsenbeck JP (2003) MrBayes3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574. Available from: http://mrbayes .sourceforge.net/. [Accessed 30 July 2013]

Rosauer DR, Biga PR, Lindell SR, Binkowski FP, Shepherd BS, Palmquist DE, Simchick CA, Goetz FW (2011) Developments of yellow perch (*Perca flavescens*) broodstocks: initial characterization of growth quality traits following grow-out of difference stocks. Aquaculture 317: 58-66

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Res 8(1): 103-106. Available from: http://genepop.curtin.edu.au/. [Accessed 30 July 2013]

Rudakova SL, Kurath G, Bochkova EV (2007) Occurrence and genetic typing of Infectious Hematopoietic Necrosis virus in Kamchatka, Russia. Dis Aquat Org 75: 1-11

Schumacher W, Frick E, Kauselmann M, Maier-Hoyle V, van der Vliet R, Babiel R (2007) Fully automated quantification of human immunodeficiency virus (HIv) type 1 RNA in human plasma by the COBAS (R) AmpliPrep/COBAS (R) TaqMan (R) system. J Clin Virol 38: 304-312

Schütze H, Mundt E, Mettenleiter TC (1999) Complete genomic sequence of Viral Hemorrhagic Septicemia virus, a fish rhabdovirus. Virus Genes 19: 59-65

Shabir GA (2003) Validation of high-performance liquid chromatography methods for pharmaceutical analysis understanding the differences and similarities between validation requirements of the US Food and Drug Administration, the US Pharmacopeia and the International Conference on Harmonization. J Chromatogr 987: 57-66

Shäperclaus W (1938) Die schädigungen der deutschen fischerei durch fischparasiten und fischkrankheiten. Fischerei-Zeitung 41: 267-270

Shapiro-Ilan DI, Fuxa JR, Lacey LA, Onstad DW, Kaya HK (2005) Definitions of pathogenicity and virulence in invertebrate pathology. J Inver Pathol 88: 1-7

Skall HF, Olesen NJ, Mellergaad S (2005) Viral Haemorrhagic Septicaemia virus in marine fish and its implications for fish farming – a review. J Fish Dis 28(9): 509-529

Smail DA (1999) Viral Haemorrhagic Septicaemia. In: PTK Woo, DW Bruno (eds.), Fish Diseases and Disorders: Viral, Bacterial and Fungal Infections, Fisheries Research Servies, Aberdeen, United Kingdom, pp. 123-147

Snow M, Bain N, Black J, Taupin V, Cunningham CO, King JA, Skall HF, Raynard RS (2004) Genetic population structure of marine Viral Haemorrhagic Septicaemia virus (VHSv). Dis Aquat Org 61: 11-21

Snow M, Cunningham CO, Melvin WT and Kurath G (1999) Analysis of the nucleoprotein gene identifies distinct lineages of Viral Haemorrhagic Septicaemia virus within the European marine environment. Virus Res 63: 35-44

Sokal RR, Rohlf FJ (1995) Biometry: 3rd edn. W. H. Freeman and Company, New York, New York

Ståhlberg A, Håkansson J, Xian X, Semb H, Kubista M (2004) Properties of the reverse transcription reaction in mRNA quantitation. Clin Chem 50: 509-515

Stepien CA, Kocher TD (1997) Molecules and morphology in studies of fish evolution. In: Kocher TD, Stepien CA (eds.), Molecular Systematics of Fishes. San Diego Academic Press, USA, pp. 1-11

Stone DM, Ferguson HW, Tyson PA, Savage J, Wood G, Dodge MJ, Woolford G, Dixon PF, Feist SW, Way K (2008) The first report of Viral Haemorrhagic Septicaemia in farmed rainbow trout, *Oncorhynchus mykiss* (Walbaum), in the United Kingdom. J Fish Dis 31: 775-784

Stone DM, Way K, Dixon F (1997) 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.). J Gen Virol 78: 1319-1326

Studer J, Janies DA (2011) Global spread and evolution of Viral Haemorrhagic Septicaemia virus. J Fish Dis 34: 741-747

Swanson P, Huang S, Holzmayer V, Bodelle P, Yamaguchi J, Brennan C, Badaro R, Brites C, Abravaya K, Devare SG, Hackett J Jr. (2006) Performance of the automated Abbott RealTime (TM) HIv-1 assay on a genetically diverse panel of specimens from Brazil. J Virol Methods 137: 184-192

Swofford DL (2003) PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Maryland. Available from: http://paup.csit.fsu.edu/. [Accessed 30 July 2013]

Tajima F (1989) Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics 123: 585-595

Takano R, Nishizawa T, Arimoto M, Muroga K (2000) Isolation of Viral Hemorrhagic Septicemia virus (VHSv) from wild Japanese flounder, *Paralichthys olivaceus*. Bull Eur Assoc Fish Pathol 20: 186-192

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10): 2731-2739. Available from: http://www.megasoftware.net/. [Accessed 30 July 2013]

Thiéry R, de Boisséson C, Jeffroy J, Castric J, de Kinkelin P, Benmansour A (2002) Phylogenetic analysis of Viral Haemorrhagic Septicaemia virus (VHSv) isolates from France (1971-1999). Dis Aquat Org 52: 29-37

Thiry M, Lecoq-Xhonneux F, Dheur I, Renard A, De Kinkelin, P (1991) Sequence of a

cDNA carrying the glycoprotein gene and part of the matrix protein *M*2 gene of Vial Haemorrhagic Septicaemia virus, a fish rhabdovirus. BBA 1090(3): 345-347

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876-4882. Available from: http://www.clustal.org/. [Accessed 30 July 2013]

Thompson TM, Batts WN, Faisal M, Bowser P, Casey JW, Phillips K, Garver KA, Winton J, Kurath G (2011) Emergence of Viral Haemorrhagic Septicaemia virus in the North American Great Lakes region is associated with low viral genetic diversity. Dis Aquat Org 96: 29-43

Thoulouze MI, Bouguyon E, Carpentier C, Bremont M (2004) Essential role of the *Nv* protein of *Novirhabdovirus* for pathogenicity in rainbow trout. J Virol 78: 4098-4107

Troyer RM, Kurath G (2003) Molecular epidemiology of Infectious Hematopoietic Necrosis virus reveals complex virus traffic and evolution within southern Idaho aquaculture. Dis Aqua Organ 55: 175-185

U.S. Department of Agriculture (USDA) and Animal Health Plant Inspection Service (APHIS) (2006) Viral Hemorrhagic Septicemia in the Great Lakes: July 2006 emerging disease notice. Available from: http://www.aphis.usda.gov/animal_ health/emergingissues/downloads/ vhsgreatlakes.pdf. [Accessed 30 July 2013]

U.S. Environmental Protection Agency (2004) Quality assurance/quality control guidance for laboratories performing PCR analyses on environmental samples. EPA 815-B-04-001

U.S. Fish and Wildlife Service and American Fisheries Society-Fish Health Section (2010) Standard procedures for aquatic animal health inspections. In: American Fisheries Society-Fish Health Section, Fish Health Section Blue Book: Suggested Procedures for the Detection and Identification of Certain Finfish and Shellfish Pathogens. American Fisheries Society-Fish Health Section, Bethesda, Maryland

U.S. Food and Drug Adminitration (2010) Draft guidance for industry and food and drug administration staff; Establishing the performance characteristics of *in vitro* diagnostic devices for the detection of *Clostridium difficile*. Center for Devices and Radiological Health. Available from: http://www.fda.gov/downloads/ MedicalDevices/ DeviceRegulationandGuidance/GuidanceDocuments/ UCM234878.pdf. [Accessed 30 July 2013]

Vogelstein B, Kinzler KW (1999) Digital PCR. Proc Natl Acad Sci 96: 9236-9241

Waknitz FW, Tyan TJ, Nash CE, Iwamoto RN, Rutter LG (2002) Review of potential impacts of Atlantic salmon culture on Puget Sound Chinook salmon and Hood Canal summer-run chum salmon evolutionarily significant units. NOAA Technical Memorandum NMFS-NWFSC-53

Walker PJ, Kongsuwan K (1999) Deduced structural model for animal rhabdovirus glycoproteins. J Gen Virol 80: 1211-1220

Washington State Department of Health (2011) Wild vs. Farmed Salmon. Division of Environmental Health, Office of Environmental Health, Safety, and Toxicology, WA. Available from: http://oregonprogress.oregonstate.edu/winter-2011/invading-species. [Accessed 30 July 2013]

Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population

structure. Evolution 38: 1358-1370

Whelan GE (2009) Viral Hemorrhagic Septicemia (VHS) briefing paper. Michigan Department of Natural Resources. Available from: http://www.michigan.gov/documents/dnr/Viral-Hemorrhagic-Septicemia-Fact-Sheet-11-9-2006_178081_7.pdf. [Accessed 23 July 2013]

Willey JC, Crawford EL, Jackson CM, Weaver DA, Hoban JC, Khuder SA, DeMuth JP (1998) Expression measurement of many genes simultaneously by quantitative RT-PCR using standardized mixtures of competitive templates. Am J Resp Cell Mol Biol 19: 6-17

Willey JC, Crawford EL, Knight CR, Warner KA, Motten CA, Herness EA, Zahorchak RJ, Graves TG (2004) Standardized RT-PCR and the standardized expression measurement center. Methods Mol Biol 258: 13-41

Winton JR, Einer-Jensen K (2002) Molecular diagnosis of infectious hematopoietic necrosis and viral hemorrhagic septicemia, In: Cunningham CO (ed.), Molecular Diagnosis of Salmonid Diseases. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 49-79

Wisconsin Aquaculture Association Inc. (2011) VHS Articles. Available from: http://www.wisconsinaquaculture.com/News_Details.cfm?NID=10&LinkType=60. [Accessed 30 July 2013]

Wisconsin Department of Natural Resources (2008) VHS fish disease found in yellow perch in Milwaukee. Available from: http://dnr.wi.gov/news/BreakingNews_Print.asp?id=864. [Accessed on 23 July 2013]

Wolf K (1988) Hemorrhagic septicemia virus, In: Wolf K (ed.) Fish Viruses and Fish Viral Diseases. Cornell University Press, Ithaca, New York, pp. 217-248

Wu X, Chi X, Wang P, Zheng D, Ding R, Li Y (2010) The evolutionary rate variation among genes of HOG-signaling pathways in yeast genomes. Biol Direct 5: 46

Yang Z (2007) PAML 4: A program package for phylogenetic analysis by maximum likelihood. Mol Biol Evol 24: 1586-1591. Available from: http://abacus.gene.ucl.ac.uk/software/paml.html. [Accessed 30 July 2013]