The Dissertation Committee for Audrey Jeanine Abrams McLean Certifies that this is the approved version of the following dissertation:

Emerging Epizootic Diseases of Amphibians and Fish: Approaches to Understanding *Ranavirus* Emergence and Spread

Committee:

David C. Cannatella, Supervisor

David M. Hillis, Co-Supervisor

James J. Bull

Lauren A. Meyers

Sara L. Sawyer

V. Gregory Chinchar

Emerging Epizootic Diseases of Amphibians and Fish: Approaches to Understanding *Ranavirus* Emergence and Spread

by

Audrey Jeanine Abrams McLean, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2013

Dedication

This dissertation is dedicated to my parents, the Reverends Robert and Carolyn Abrams. I am grateful for all of your encouragement, support, and love throughout the years. Thank you for taking time to read to me about dinosaurs, for sending me to "God Camp" to learn more about evolution, and for allowing me to conduct science experiments in the house (sorry about the fruit flies). I love you.

Acknowledgements

I would like to thank the Cannatella, Hillis, Bull, Sawyer, and Chinchar laboratory groups for their support and helpful comments and suggestions. I would also like to acknowledge funding support from the following sources: NSF IGERT Fellowship in Computational Phylogenetics and Applications to Biology Program, NSF Graduate Research Fellowship Program, Texas Ecolabs Research Grant Program, Explorer's Club Research Grant Program, and the UT-Austin EEB Research Grant Program.

Emerging Epizootic Diseases of Amphibians and Fish: Approaches to Understanding *Ranavirus* Emergence and Spread

Audrey Jeanine Abrams McLean, Ph.D.

The University of Texas at Austin, 2013

Supervisor: David C. Cannatella and David M. Hillis

Ranaviruses are large dsDNA viruses that are considered emerging pathogens, and they are known to cause mortality events in amphibian and fish populations. This research utilizes experimental and genomic data to elucidate the mechanisms driving the evolution and spread of ranaviruses, with a focus on host switching within the genus. In Chapter 1, we utilize virus challenge assays to examine potential transfer of ranaviruses between cultured juvenile largemouth bass (*M. salmoides*) and bullfrog tadpoles (*Rana catesbeiana*). Additionally, a commonly used antiparasitic treatment containing malachite green and formalin (MGF) was utilized to suppress the immune system of largemouth bass to assess the susceptibility of immunocompromised fish to ranaviruses. The results indicate that tadpoles are not susceptible to Largemouth Bass Virus (LMBV), but that bass are susceptible to ranaviruses isolated from amphibians. Furthermore, immunocompromised fish were more susceptible to both LMBV and FV3 infections than immunocompetent fish. In Chapter 2, we used eight sequenced ranavirus genomes and two selection-detection methods (site-based and branch-based) to identify genes that exhibited signatures of positive selection, potentially due to the selective pressures at play during host switching. We found evidence of positive selection acting on four genes via the site-based method, three of which are newly-acquired genes unique to ranavirus

genomes. Our results suggest that the group of newly acquired genes in the ranavirus genome may have undergone recent adaptive changes that have facilitated interspecies and interclass host switching. In Chapter 3, we annotated and analyzed the nearly complete genomic sequence of LMBV to determine its taxonomic classification. The available genomic content and phylogenetic evidence suggests that LMBV is more closely related to amphibian-like ranaviruses (ALRVs) than grouper ranaviruses, and this is further supported by greater genomic collinearity between LMBV and ALRVs. This data suggests that the classification of LMBV as a ranavirus is warranted. The results presented here will help to clarify the taxonomic relationships of ranaviruses, and will also be useful in developing management strategies to limit interspecific and intraspecific viral spread. The information garnered from this research will have far-reaching implications in studies of amphibian conservation, disease evolution, and virology.

Table of Contents

List of Tables

List of Figures

Chapter 1: Increased susceptibility of largemouth bass (*Micropterus salmoides*) to ranaviruses when exposed to a commonly used disease treatment containing malachite green and formalin

ABSTRACT

Ranaviruses are large dsDNA viruses that cause recurrent mortality events in wild and commercial amphibian and fish populations. Ranaviruses are considered emerging pathogens, and researchers have developed several hypotheses to explain the emergent nature of the viruses including: 1) introduction into naïve host populations via amphibian or fish vectors; 2) geographic spread through the bait, pet, and food industries; and 3) adverse effects on the immune system due to extrinsic immunocompromising agents. Here, we utilize virus challenge assays to examine the role that each mechanism may play in the transfer of ranaviruses between cultured juvenile largemouth bass (*Micropterus salmoides*) and bullfrog tadpoles (*Rana catesbeiana*), which can cohabit in aquaculture settings. Additionally, a commonly used antiparasitic treatment containing malachite green and formalin (MGF) was utilized to suppress the immune system of largemouth bass to assess the susceptibility of immunocompromised fish to ranaviruses. The results indicate that bullfrog tadpoles are not susceptible to a ranavirus isolated from fish, but that largemouth bass are susceptible to ranaviruses isolated from amphibians. Immunocompromised fish were more susceptible to ranavirus infections than immunocompetent fish. This suggests that juvenile largemouth bass may act as a reservoir for ranaviruses that commonly infect bullfrogs, and that the susceptibility of the fish may be heightened after exposure to commonly used antiparasitic treatments. Considering the extensive introduction of largemouth bass worldwide, precautions should

be taken to limit contact between cultured largemouth bass and wild amphibian populations. Furthermore, management strategies should be developed to the prevent introduction of infected fish to naïve habitats where native amphibian populations may be at risk.

INTRODUCTION

The genus *Ranavirus* includes large dsDNA viruses (150 nm) that cause recurrent mortality events in wild and commercial amphibian and fish populations, and some strains are known pathogens of reptiles (Chinchar 2002; Whittington *et al*. 2010; Miller *et al*. 2011). Ranaviruses are considered emerging due to the relatively recent increase in incidence, geography, and host range (Daszak *et al*. 2000); isolates have been detected on every continent with the exception of Africa and Antarctica. Most ranaviruses cause systemic infections involving the liver, kidneys, and gastrointestinal tract, which may result in hemorrhages of these organs (Williams *et al*. 2005; Whittington *et al*. 2010). However, asymptomatic infections are common, and several *in vivo* studies have identified susceptible hosts that have no documentation of infection in wild populations (Cullen & Owens 2002; Schock *et al*. 2008; Bayley *et al*. 2013; Becker *et al*. 2013). This suggests that numerous species may have the potential to act as viral reservoirs and vectors.

Several hypothesis have been put forth to explain the emergent nature of ranaviruses. Researchers have proposed that introduction into naïve host populations via amphibian or fish vectors may facilitate the spread to new species. For instance, Mao *et al*. (1999a) isolated identical ranavirus specimens from sympatric red-legged frogs (*Rana aurora*) and three-spine stickleback fish (*Gasterosteus aculeatus*), thus suggesting the

possible role of one of the species as a vector or reservoir for the virus. Moreover, the bait, pet, and food industries are considered major contributors to the spread of ranaviruses to naïve localities. This hypothesis is supported by phylogenetic and physical evidence that ranaviruses isolated from tiger salamanders (*Ambystoma tigrinum*) are spread throughout the western United States via the bait trade (Jancovich *et al*. 2005; Picco & Collins 2008). Lastly, adverse effects on the immune system via extrinsic immunocompromising agents or outbreeding depression have also been proposed as mechanisms that enable epidemic transmission of ranaviruses. With regard to the role of immunosuppression in amphibians, researchers have hypothesized that single or multiple combinations of stressful environmental changes may alter the neuroendocrine system, resulting in immunosuppression (Carey *et al*.1999). For instance, researchers determined that exposure to the common pesticides atrazine and chlorpyrifos increased the susceptibility of tiger salamanders to ranaviruses (Forson & Storfer 2006; Kerby & Storfer 2009.) Thus, immunosuppression could leave individuals more susceptible to ranavirus infections, or facilitate pathogen host switching. Furthermore, a study of outbreeding depression in largemouth bass (*Micropterus salmoides*) determined that hybrid individuals were more susceptible to a ranavirus strain than their wild-type parents (Goldberg *et al*. 2005). Although evidence supports each hypothesis, the emergent nature of ranaviruses is most likely due to the synergistic effects of the mechanisms discussed above.

Bullfrogs (*Rana catesbeiana*) and largemouth bass (*M. salmoides*) represent two abundant and commonly introduced species, which are known hosts of the ranaviruses Frog Virus 3 (FV3) and Largemouth Bass Virus (LMBV), respectively. Together, the two species represent an ideal system to study interspecific ranavirus reservoirs and the potential spread through anthropogenic avenues. Both species are native to the eastern

United States, and are regularly farmed on a global scale for food, scientific, and recreational purposes (Schloegal *et al*. 2009; Cooke *et al*. 2002). Bullfrogs have been introduced to over 40 countries and four continents, and ranavirus infections have been detected in both cultured and wild populations (Green *et al*. 2002; Hanselmann *et al*. 2004; Gray *et al*. 2007; Miller *et al*. 2007). Similarly, the largemouth bass is considered an alien species in nearly 80 countries (Global Invasive Species Database 2013), and LMBV infections have been detected in wild and cultured populations across the southeastern United States (Plumb *et al*. 1999; Hanson *et al*. 2001; Maceina & Grizzle 2006; Neal *et al*. 2009; Southard *et al*. 2009; Blazer *et al*. 2010). The open pond designs of many aquaculture facilities often result in habitation by wild frogs and tadpoles (Corse & Metter 1980; Carmichael & Tomasso 1983; Kloskowski 2010). Bullfrog tadpoles inhabit artificial fishponds where fingerling largemouth bass are held (Abrams personal observation), which permits ranaviruses to spread between the two species. Picco *et al*. (2010) showed that asymptomatic largemouth bass were susceptible to a ranavirus isolated from salamanders. Therefore, it is possible that upon exposure to infected amphibians, cultured largemouth bass could become infected. Many cultured largemouth bass are relocated to stock private and public bodies of water that are open to wild amphibian habitats. If infected while in captivity, the relocated fish could potentially transfer ranaviruses to naïve amphibian habitats.

The possible spread of ranaviruses between cohabitating bullfrog tadpoles and fingerling largemouth bass could be exacerbated by the stressful conditions associated with general fish husbandry, collection, and transport. Fish eggs and fish stocks are regularly treated for bacterial, fungal, and protozoan infections with antiparasitic solutions containing malachite green and formalin (Meinelt *et al*. 2001; Srivastava *et al*. 2004; Sudova *et al*. 2007). Malachite green is an organic dye, and formalin is an aqueous

solution of the organic compound formaldehyde. Exposure to antiparasitic solutions containing malachite green and formalin (hereafter referred to as MGF treatments) can elicit significant stress responses, and malachite green is a known immunosuppresent in fish (Culp and Beland 1996; Yildiz and Pulatsu 1999; Kodama *et al*. 2004, Silveira-Coffigny *et al*. 2004; Yonar and Yonar 2010). The reduced capacity of the immune system may result in increased viral proliferation; therefore, fish cohabiting with diseased wild amphibians may be more susceptible to ranaviruses after exposure to commonly used antiparasitic treatments.

Although several viral challenge studies have reported the susceptibility of fish to ranaviruses isolated from amphibians (Moody & Owens 1994; Jancovich *et al*. 2001; Gobbo *et al*. 2010; Bang Jensen *et al*. 2009; Picco & Collins 2010), only one study has examined the susceptibility of amphibians to ranaviruses isolated from fish (Bayley *et al*. 2013). Here, we present data from viral challenge assays used to determine the susceptibility of bullfrog tadpoles to LMBV and the susceptibility of juvenile largemouth bass to FV3 and a virus isolated from salamanders (Spotted Salamander Maine Virus, SSMe). We subsequently examined variation in the susceptibility of largemouth bass to LMBV and FV3 under conditions where fish were regularly exposed to immunocompromising MGF treatments and those where they were not. The results indicate that bullfrog tadpoles are not susceptible to LMBV, but that largemouth bass are susceptible to FV3 and SSMe. Furthermore, fish challenged under immunocompromising conditions were more susceptible to both LMBV and FV3 infections than untreated fish. However, the difference in susceptibility between treatments was more pronounced in fish exposed to FV3. This data suggests that juvenile largemouth bass may act as a reservoir for FV3-like viruses that commonly infect

bullfrogs, and that the susceptibility of the fish may be heightened after exposure to commonly used antiparasitic treatments containing malachite green and formalin.

MATERIALS AND METHODS

Two viral challenge assays were performed. The purpose of Challenge I was to determine the susceptibility of bullfrog tadpoles (*R. catesbeiana*) to LMBV (isolated from fish), and juvenile largemouth bass (*M. salmoides*) to FV3 and SSMe (isolated from amphibians). Although FV3 and LMBV are known pathogens of bullfrogs and largemouth bass respectively, experimental evidence has shown variation in both the susceptibility of bullfrogs to FV3 and largemouth bass to LMBV (Woodland *et al*. 2002; Goldberg *et al*. 2003; Miler *et al*. 2007; Hoverman *et al*. 2011). Thus, trials using challenging each host with its respective pathogen were used to determine the sensitivity of the specific populations used in the study. The purpose of Challenge II was to determine if largemouth bass exposed to an immunocompromising antiparasitic treatment containing malachite green and formalin (MGF) exhibited heightened susceptibility to FV3 and LMBV when compared to those that did not receive the MGF treatment. A subset of both tadpoles and fish were humanely euthanized and screened for ranavirus DNA before the start of each trial, and no evidence of infection was detected. The husbandry and experimental conditions for both challenge assays are described below.

Animal Husbandry: Challenge I- Tadpole and Fish Susceptibility

Tadpoles

Field-collected *R. catesbeiana* tadpoles (1-3 cm) were obtained from Carolina Biological Supply Company (Burlington, NC). Tadpoles were housed in 58 L aquaria

that contained gravel and treated tap water. Tadpoles were acclimated in the laboratory for at least two weeks prior to virus exposure, and were held under a 12-hr. light: 12-hr. dark cycle. During the acclimation and experimental periods, tadpoles were fed tadpole pellets (Zoo Med Laboratories, Inc.) *ad libitum* once per day, and water was changed weekly.

Fish

Juvenile *M. salmoides* (5-8 cm) were obtained from a commercial fish farm (Fittstown, Oklahoma). Fish were housed in 38 L, 58 L, 78 L, or 246 L circulating aquaria during an acclimation period for at least two weeks prior to virus exposure and were held under a 12-hr. light: 12-hr. dark cycle. During the experimental period, all fish were housed in in 58 L circulating aquaria that contained gravel and treated tap water. During the acclimation and experimental periods, fish were fed cichlid pellets (OmegaSea, Ltd.) *ad libitium* once per day, and water was changed weekly. A fungal infection was detected in all experimental tanks on Day 9 of the experiment, and fish were treated daily with an MGF solution until visible signs of infection were eliminated (Day 15). Each dose of the MGF solution contained a malachite green concentration of 0.05 ppm and a formalin concentration of 15 ppm after the solution was added to the tank. A partial water change (approximately 25%) was performed prior to the addition of the MGF solution, and carbon filters were removed. During treatment, fish were exposed to the MGF solution for approximately 7 hours. Carbon filters were replaced to absorb malachite green and formalin (Aitcheson *et al*. 2000), and the water was filtered for approximately 17 hours. Full water changes were completed on a weekly basis.

Animal Husbandry: Challenge II- Susceptibility of Immunocompromised Fish

Juvenile largemouth bass were obtained, acclimated, and maintained as described above. During the first week of acclimation, all fish were treated every other day with an MGF solution to eliminate fungal infections and other parasites. Each dose of the MGF solution contained a malachite green concentration of 0.05 ppm and a formalin concentration of 15 ppm. A partial water change (approximately 25%) was performed prior to the addition of the MGF solution, and carbon filters were removed. During treatment, fish were exposed to MGF solution for approximately 7 hours. Carbon filters were replaced to absorb malachite green and formalin, and the water was filtered for approximately 17 hours. Fish were fed cichlid pellets (OmegaSea, Ltd.) once per day, and full water changes were completed each week.

During the experimental period, all fish were housed in in 58 L circulating aquaria that contained gravel and treated tap water, and were held under a 12-hr. light: 12-hr. dark cycle. Fish were fed cichlid pellets once per day, and full water changes were completed on a weekly basis.

Virus Stocks

Frog Virus 3 (FV3), Spotted Salamander Maine Virus (SSMe), and Largemouth Bass Virus (LMBV) virus stocks that were grown in fathead minnow cell culture were obtained from Dr. V. Gregory Chinchar (University of Mississippi Medical Center). The titers of the viral stocks used in Challenge I were FV3 at 2.3 x 10^7 pfu/ml, SSMe at 4.8 x 10^7 pfu/ml, and LMBV at 2.0 x 10^7 pfu/ml. The titers of the viral stocks used in Challenge II were FV3 at 5.3 x 10^6 pfu/ml and LMBV at 3.0 x 10^5 $TCID_{50}$/ml. Virus

stocks were stored at -80 °C until utilized to avoid repeat freeze/thaw cycles that could result in a lower viral concentration.

Viral Challenges

Challenge 1: Fish and Tadpole Susceptibility

Tadpoles were housed in three 58 L circulating aquaria (mean temperature = 20.9 °C, min-max temperature = 20.0-21.6 °C) that contained gravel and treated tap water, and 23 individuals were held in each tank (Table 1.1). Sterile 27-gauge needles were used to inject each animal intraperitoneally with 0.1 mL of FV3, LMBV, or a sham injection of Eagle's minimum essential medium with Hank's salts. Tanks were monitored daily to identify and remove dead tadpoles, which were then stored at -80°C. Moribund tadpoles were euthanized via an overdose of tricaine-methylsulfate, and subsequently stored at -80°C until processing.

Fish were housed in four 58 L circulating aquaria (mean temperature = 21.0 °C, min-max temperature = 19.7-21.9 °C) that contained gravel and treated tap water, and 13 individuals were held in each tank (Table 1.1). Fish were sedated with tricaine-methylsulfate prior to receiving a 0.1 mL intraperitoneal injection of FV3, SSMe, LMBV, or a sham injection of Eagle's minimum essential medium with Hank's salts. Tanks were monitored daily to identify and remove dead fish, which were then stored at -80°C. Moribund fish were euthanized via an overdose of tricaine-methylsulfate, and subsequently stored at -80°C until processing.

Challenge II: Susceptibility of Immunocompromised Fish

Virus Challenge Only (No MGF treatment): Fish were housed in 58 L circulating aquaria (mean temperature = 20.4 °C, min-max temperature = 20.0-22.1 °C) that contained gravel and treated tap water. Three treatments were used (FV3, LMBV, and Control), and three replicate tanks with nine individuals each were used for each treatment (Table 1.1). Fish were sedated with tricaine-methylsulfate prior to receiving a 0.1 mL intraperitoneal injection of FV3, LMBV, or a sham injection of Eagle's minimum essential medium with Hank's salts. Tanks were monitored daily to identify and remove dead fish, which were then stored at -80°C. Moribund fish were euthanized via an overdose of tricaine-methylsulfate, and subsequently stored at -80°C until processing.

Virus Challenge Combined with MGF Treatment: Fish were housed in 58 L circulating aquaria (mean temperature = 20.0 °C, min-max temperature = 19.4-22.2 °C) that contained gravel and treated tap water. Three treatments were used (FV3, LMBV, and Control), and four replicate tanks with nine individuals each were used for each treatment (Table 1.1). Fish were sedated with tricaine-methylsulfate prior to receiving a 0.1 mL intraperitoneal injection of FV3, LMBV, or a sham injection of Eagle's minimum essential medium with Hank's salts. All tanks were treated daily with an MGF solution containing a malachite green concentration of 0.05 ppm and a formalin concentration of 15 ppm. A partial water change (approximately 25%) was performed prior to the addition of the MGF solution, and carbon filters were removed. During treatment, fish were exposed to MGF solution for approximately 7 hours. Carbon filters were replaced to absorb malachite green and formalin, and the water was filtered for approximately 17 hours. Tanks were monitored daily to identify and remove dead fish, which were then

stored at -80°C. Moribund fish were euthanized via an overdose of tricaine-methylsulfate, and subsequently stored at -80°C until processing.

Viral DNA Isolation, Polymerase Chain Reaction (PCR), and Sequencing

Prior to dissection, the weight and length of each tadpole (total length) and fish (standard length) was measured and recorded. A liver sample was taken from each individual tadpole using sterile techniques. Liver, spleen, and trunk kidney samples were dissected from each fish using sterile techniques. Cross-contamination was minimized by processing individuals from each treatment separately. DNA was isolated from individual tissue samples using the standard protocol from the DNeasy Blood and Tissue Kit (Qiagen). PCR was performed using 25 μL reactions containing 0.2 μL AmpliTaq Gold DNA Polymerase (Invitrogen), 2.5 μL 10X buffer, 1.5 μL $MgCl_2$, 2.5 μL dNTP mix, 1 μL of each 20 μM primer, 2 μL template DNA, and 14.3 μL H_2O. The hot-start thermocycler conditions were as follows: 95 °C for 9 min, 40 cycles of 94 °C for 30 sec., 50 °C for 30 sec., 72 °C for 45 sec, and 72 °C for 7 min. Primers described in Mao *et al*. (1997) were used to amplify a 500 bp fragment of the ranavirus major capsid protein (MCP) gene. The sequences of positive PCR products were obtained via Sanger sequencing, and the resulting sequence was compared to nucleotide sequences stored in the National Center for Biotechnology Information database using the BLASTN similarity search algorithm (http://blast.ncbi.nlm.nih.gov) to confirm the presence of ranavirus DNA.

Statistical Analyses

To determine if there were significant differences among the survivorship curves for each treatment, Mantel-Cox log-rank tests (α= 0.05, one-tailed test) were conducted

for each pair of curves. Moreover, for Challenge II, the cumulative survivorship curves for pairs of virus only and virus combined with MGF treatment curves were tested. To assess whether the weight and length of the fish tested in Challenge II significantly affected death or survivorship, the median and mean values for both variables, for each treatment group, were compared using Kruskal-Wallis (median) and ANOVA (mean) tests.

RESULTS

Challenge 1: Tadpole and Fish Susceptibility

No evidence of bullfrog tadpole susceptibility to LMBV

In the bullfrog tadpole susceptibility trial, the highest mortality was seen in the FV3 treatment (Figure 1.1). Although mortality was present in all treatments, no viral DNA was detected in sham- or LMBV-injected animals (Table 1.2). In contrast, FV3 DNA was detected in treated tadpoles by PCR and confirmed by sequencing a fragment of the MCP gene. Comparisons of the survivorship curves indicate significant differences between the control and FV3 treatments ($p=0.0038$) and the FV3 and LMBV treatments ($p= 0.0396$). However, no significant difference was found between the control and LMBV treatments. These results indicate that the tadpoles used in this study are susceptible to FV3, but not LMBV.

FV3, SSMe, and LMBV DNA isolated from juvenile largemouth bass

In the largemouth bass susceptibility trial, mortality was observed in all treatments (Figure 1.2). However, the presence and treatment of a visible fungal infection with the MGF solution confounds the interpretation of the results. Prior to the

visual evidence and treatment of the fungal infection, the survivorship curves show early instances of death in the FV3 treatment; however, the fungal infection could have been present, but not apparent. During the experiment (including the infection and treatment period) all tanks received the same treatment regimen, so comparisons of the survivorship curves to detect significant differences were performed. Significant differences were only found between control and FV3 treatments ($p = 0.0233$) and FV3 and LMBV treatments ($p = 0.0077$). No viral DNA was isolated from the animals in the control treatment, but viral DNA was isolated from each viral treatment (Table 1.2). The cause of death in the fish tested cannot be attributed to the viral infection because of the confounding effects of the fungal infection and MGF treatment. However, the isolation and verification of viral DNA from fish (until Day 15 of the experiment), strongly suggests that the fish tested were susceptible to FV3, SSMe, and LMBV.

Challenge II: Susceptibility of Immunocompromised Fish

The isolation of ranavirus DNA from the fish tested in Challenge I supports the hypothesis that the largemouth bass population tested is susceptible to LMBV, FV3, and SSMe. Therefore, we conducted further experiments to compare the susceptibility of juvenile largemouth bass (from the same population) to LMBV and FV3 under conditions where the fish were immunocompromised (via daily exposure to MGF treatments) and immunocompetent (no exposure to MGF treatments). Although the isolation of DNA from infected fish was an ideal verification method for positive ranavirus infection, DNA extraction attempts yielded degraded genomic DNA that could not be used for viral DNA detection in subsequent applications.

Virus challenge only (no MGF treatment) indicates moderate susceptibility to FV3 in immunocompetent fish

Mortality was detected in trials where fish were exposed to FV3 and LMBV in the absence of daily MGF treatments. The cumulative survivorship of fish exposed to LMBV was significantly lower than that of fish injected with a sham ($p = <0.0001$), and marginally significantly lower than those injected with FV3 ($p = 0.457$) (Figure 1.3A). DNA extraction attempts yielded degraded DNA, so verification of viral DNA was not possible. However, previous evidence of largemouth bass susceptibility to FV3 and LMBV, in combination with these results, suggests that the immunocompetent fish are susceptible to LMBV and moderately susceptible to FV3 in the absence of daily MGF treatments.

Virus challenge in combination with MGF treatment indicates susceptibility to FV3 and LMBV in immunocompromised fish

In the presence of daily MGF treatments, mortality was observed in control, LMBV, and FV3 trials (Figure 1.3B). Although the cumulative survival seen in the control group was significantly different from that of the immunocompetent control group ($p = 0.0019$), it differed significantly from the survivorship of FV3 ($p = 0.0001$) and LMBV ($p = 0.0002$) (Table 1.3). The cumulative survivorship of fish exposed to FV3 was not significantly different from that of those exposed to LMBV. These results suggest that the daily MGF treatment alone contributes to mortality in largemouth bass, but that the combination of the MGF treatment with viral infection significantly increases mortality. DNA extraction attempts yielded degraded DNA, so verification of viral DNA was not possible.

Significant differences between cumulative survival of immunocompetent and immunocompromised fish suggest heightened ranavirus susceptibility in immunocompromised fish

Direct comparisons of the cumulative survival of immunocompetent fish (no daily MGF treatments) to those of immunocompromised fish (daily MGF treatments) yield interesting results. For the LMBV trials, the cumulative survivorship of immunocompromised fish differed significantly from the immunocompetent fish ($p = 0.0031$, Figure 1.3C). In both trials, mortality was observed throughout the length of the experiment. In contrast, the FV3 trials indicate early mortality (until Day 9) in the immunocompetent fish and continuous mortality in the immunocompromised fish. Comparisons of FV3 survivorship curves indicated a significant difference between the two trials ($p = <0.0001$, Figure 1.3D). This suggests that daily exposure to the MGF treatment significantly increases the susceptibility of juvenile largemouth bass to *Ranavirus* infection, and that this effect is more pronounced in the FV3 treatments.

DISCUSSION

The data presented here represent the first experiment used to examine the susceptibility of an amphibian species to LMBV, and the first endeavor to determine the effects of a commonly used antiparasitic treatment (MGF) on the susceptibility of largemouth bass to ranaviruses. Although LMBV is a multihost pathogen known to infect at least seventeen species of fish (Plumb & Zilberg 1999; Iwanowicz *et al*. 2013), its inability to infect bullfrog tadpoles will likely be repeated when other amphibian species are tested. Phylogenetic analyses of ranaviruses place LMBV between the grouper-like ranaviruses and the amphibian-like ranaviruses (including FV3 and SSMe), which

indicates that LMBV evolved before the initial host switch into amphibians (Mao *et al*. 1999b; Mavian *et al*. 2012; Abrams unpublished data). LMBV may simply lack the genetic mechanisms needed to bypass the amphibian immune system.

Unlike LMBV, FV3-like viruses are known to infect fish, amphibian, and reptile hosts (Mao *et al*. 1999a; Docherty *et al*. 2003; Huang *et al*. 2009; Miller *et al*. 2011); thus, suggesting that these viruses not only evolved the ability to infect amphibians, but also maintained the genetic capabilities to infect fish. The results of the Challenge I experiments indicate the ability of FV3-like viruses to infect juvenile bass, but any symptoms associated with the infection may have been masked by the concurrent fungal infection. A previous investigation into the susceptibility of largemouth bass to a *Ranavirus* isolated from salamander larvae (commonly used as bait) discovered evidence of asymptomatic infection (Picco *et al*. 2010). Picco *et al*. (2010) suggested that largemouth bass may act as a reservoir or vector for this particular strain, and the same may be true with regard to the FV3-like viruses (FV3 and SSMe) that were tested in this study. Largemouth bass regularly come into contact with potentially infected salamander larvae (eaten as fishing bait) and bullfrog tadpoles (cohabitants of artificial ponds), which could provide opportunities for amphibian ranaviruses to adapt to a largemouth bass host. Likewise, in the wild, largemouth bass are known to consume several salamander and frog species that are naturally occurring *Ranavirus* hosts (Hodgson & Hansen 2005). However, in both challenge studies, the fish were infected via injections, which is obviously not a natural infection route. Future studies should test the susceptibility of largemouth bass to FV3-like viruses (and the transmission of these viruses from fish back into amphibians) via natural routes such as exposure to virus-infected water and cohabitation with infected amphibians.

Results from Challenge II experiments where juvenile largemouth bass were exposed to a common antiparasitic treatment (MGF) suggest that exposure inhibits the ability of the fish to combat ranavirus infections (Figure 1.3). Significant differences between the control groups of immunocompromised and immunocompetent fish indicated that exposure to MGF alone significantly affected fish survival. Furthermore, the mortality observed in LMBV and FV3 treatments suggests that exposure to MGF heightens susceptibility to both ranaviruses, but that the effect is more pronounced in FV3 treatments (Figure 1.3C-D). Viral reservoirs are often more resistant to the pathogen than the targeted host species (Brunner *et* al. 2004), and this characteristic is reflected in the significantly higher survival rate observed in immunocompetent fish exposed to FV3 than that seen in immunocompromised fish (Figure 1.4D). The relatively high survival of immunocompetent fish exposed to FV3 suggests that they are only moderately affected by the virus. Consequently, juvenile largemouth bass that are exposed to MGF treatments are more likely to contract an FV3-like virus than their immunocompetent counterparts, and may possibly transmit the virus to new habitats and species if relocated.

The heightened susceptibility of immunocompromised fish reared in contact with potentially infected tadpoles is not only a concern with regard to the use of MGF treatments, but also other stress factors that can negatively affect fish immune function. Artificial stress factors that are common in aquaculture such as chemical exposure, poor handling, and confinement are associated with suppressed immune systems in fish, which may predispose fish in these environments to infectious diseases (Ellaesser & Clem 1986; Bly *et al*. 1997; Harris & Bird 2000). In fact, the confinement of largemouth bass during fishing tournaments corresponds to increased prevalence of LMBV in largemouth bass populations (Schramm *et al*. 2006). Further investigations into the susceptibility of

largemouth bass to FV3-like viruses should evaluate the effects of artificial stressors on fish susceptibility.

The use of non-chemical artificial stressors may prove more useful in studies of fish susceptibility to ranaviruses than the chemicals utilized here. The exposure of fish to MGF treatments resulted in the extensive degradation of genomic DNA, so detection of ranavirus DNA in Challenge II fish was not possible. The digestion of tissues extracted from fish that received daily MGF treatments resulted in greenish-blue homogenates due to the high concentration of malachite green that is often stored in the liver, kidneys, and spleen (Culp & Beland 1996). It is also likely that formalin was also concentrated in these organs, and both malachite green and formalin are known to degrade nucleic acids (Ma & Harris 1988; Culp & Beland 1996). The inability to isolate DNA from the tissues of fish that were not regularly exposed to MGF treatments was unexpected considering DNA was extracted from Challenge I fish, which were exposed to the treatment for seven consecutive days. The most feasible explanation for this discrepancy is the length of time that the fish were frozen. Fish from Challenge I were processed within a week after the end of the experiment, while fish from Challenge II (no MGF exposure) were frozen for over a month before processing. Repeat extractions from Challenge I fish more than a year after the original extraction also yielded highly degraded DNA. Several chemical processes occur during freezing, including the oxidation of leucomalachite green to malachite green and interactions between aldehydes and fish proteins (Culp & Beland 1996; Chaijan *et al*. 2007); therefore, it is possible that the length of the freezing process may have resulted in conditions that enhanced the degradation of fish tissues and DNA. Based on these results, the use of non-chemical immunosuppressant factors would greatly increase our ability to isolate and quantify ranavirus DNA in challenged fish.

Evidence of emerging infectious diseases associated with introduced North American aquatic species is compelling, with several bacterial, fungal, viral (including ranaviruses), and protozoan pathogens linked to European disease emergence alone (Peeler *et al*. 2011). Furthermore, the recent increase in the scale of aquaculture production has facilitated the emergence of diseases through pathogen exchange with wild populations (Murray & Peeler 2005). The largemouth bass has been named one of the world's 100 worst invaders, and it is considered an alien species in over 80 countries (Global Invasive Species Database 2013). Consequently, the susceptibility of juvenile largemouth bass to FV3-like viruses is of particular concern when the introduction of the species to new habitats is considered. The open pond designs utilized by many aquaculture facilities worldwide may inevitably allow the contact between cultured fish and wild amphibian populations. The introduction of infected stocks to naïve localities may facilitate ranavirus emergence and host switching (Keisecker *et al*. 2001; Murray *et al*. 2005). The spread of ranaviruses to new amphibian populations has been linked to anthropogenic activities including the bait, pet, and food trade, and largemouth bass have already been considered potential ranavirus vectors (Picco *et al*. 2010). The added complication of heightened susceptibility to ranaviruses when exposed to commonly used antiparasitic treatments (MGF) signifies the urgent need to better understand the role that cultured largemouth bass may play in the spread of ranaviruses to amphibian populations. Considering the extensive introduction of largemouth bass worldwide, precautions should be taken to limit contact between cultured largemouth bass and wild amphibian populations. Furthermore, management strategies (i.e. selective screening for ranaviruses) should be developed to prevent the introduction of infected fish to naïve habitats where native amphibian populations may be at risk.

Experiment	Virus	Number of Tanks	Number of Individuals per Tank	Total Number of Fish Injected
Challenge I: Tadpole Susceptibility				
	FV3	1	23	23
	LMBV	1	23	23
	Control	1	23	23
Challenge I: Fish Susceptibility				
	FV3	1	13	13
	SSMe	1	13	13
	LMBV	1	13	13
	Control	1	13	13
Challenge II: Virus Challenge Only (No MGF treatment)				
	FV3	3	9	27
	LMBV	3	9	27
	Control*	3	8-9	26
Challenge II: Virus Challenge Combined with MGF Treatment				
	FV3	4	9	36
	LMBV	4	9	36
	Control	4	9	36

Table 1.1: Experimental setup for Challenge I and Challenge II.

* One control tank housed eight fish instead of nine.

Test Organism	Ranavirus Isolate	Weeks Post Infection	Number of Animals Tested*	Number Positive Animals by PCR†	Proportion Ranavirus-Positive
R. catesbeiana	FV3	Week 1	9	5	0.56
		Week 2	2	1	0.50
		Week 3	12	7	0.58
		Total	**23**	**13**	**0.56**
	LMBV	Week 1	1	0	0.00
		Week 2	2	0	0.00
		Week 3	20	0	0.00
		Total	**23**	**0**	**0.00**
M. salmoides	FV3	Week 1	5	4	0.80
		Week 2	6	4	0.67
		Week 3	2	2	1.00
		Total	**13**	**10**	**0.77**
	SSMe	Week 1	3	3	1.00
		Week 2	4	4	1.00
		Week 3	6	4	0.67
		Total	**13**	**11**	**0.84**
	LMBV	Week 1	0	0	0.00
		Week 2	7	7	1.00
		Week 3	6	1	0.17
		Total	**13**	**8**	**0.62**

Table 1.2: Challenge I: Detection of viral DNA in bullfrog tadpoles and largemouth bass challenged with various ranaviruses.

* Number of animals dead or euthanized during the specified week.
† Amplification of ranavirus major capsid protein (MCP).

Survival Curves Tested	p-value for Viral Challenge Only	p-value for Viral Challenge Combined with MGF treatment
Control vs. LMBV vs. FV3	<0.0001	0.0001
Control vs. LMBV	<0.0001	0.0002
Control vs. FV3	0.0433	0.0001
LMBV vs. FV3	0.0157	0.8630

Table 1.3: Challenge II: Results of survival curve comparisons based on Mantel-Cox log-rank test.

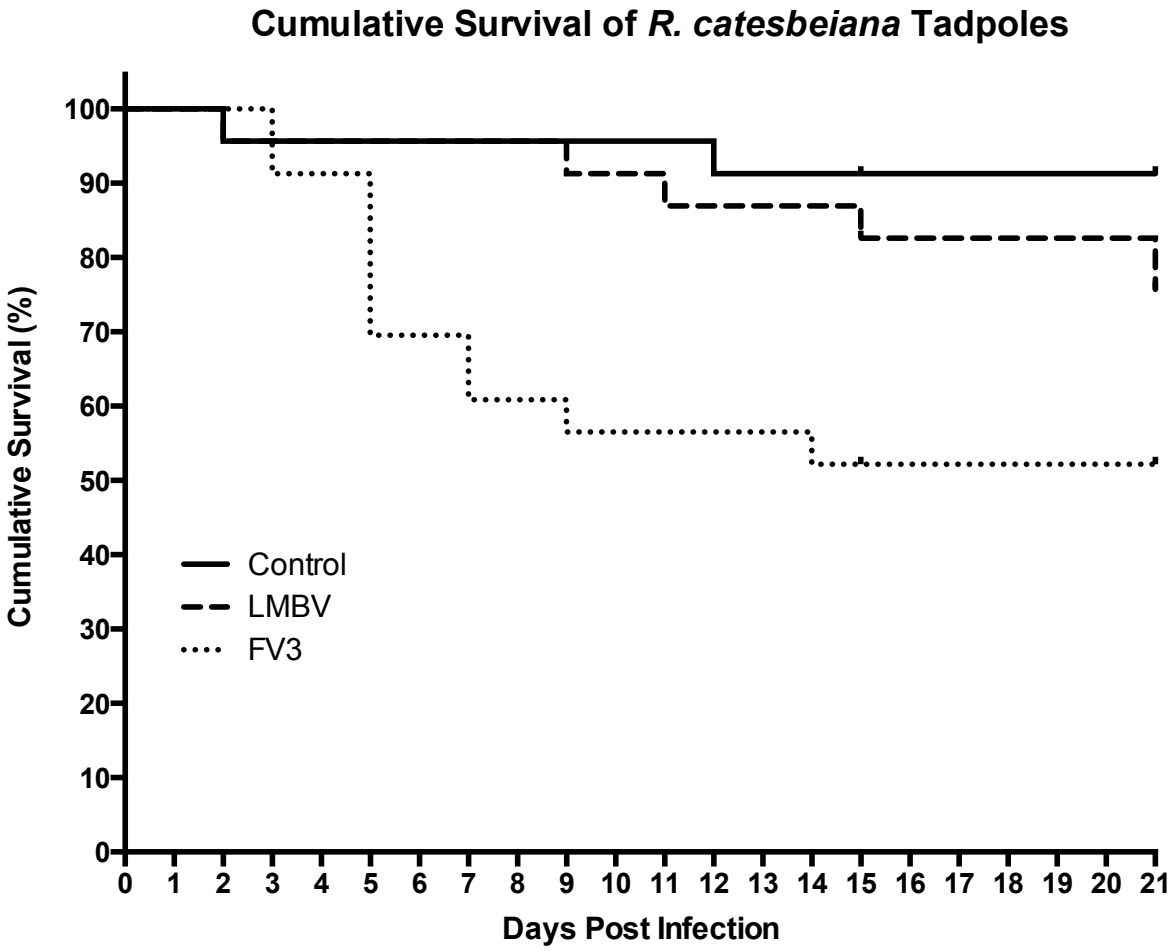

Figure 1.1: Challenge I: Percent cumulative survival of *R. catesbeiana* tadpoles infected with FV3 and LMBV. Cumulative survival for the negative controls is shown as a solid line. Significant differences in the survival curves were calculated using a Mantel-Cox log-rank test. No significant difference was found between the control and LMBV treatments ($p = 0.2287$). Significant differences were found between the control and FV3 treatments ($p = 0.0038$) and the FV3 and LMBV treatments ($p = 0.0396$).

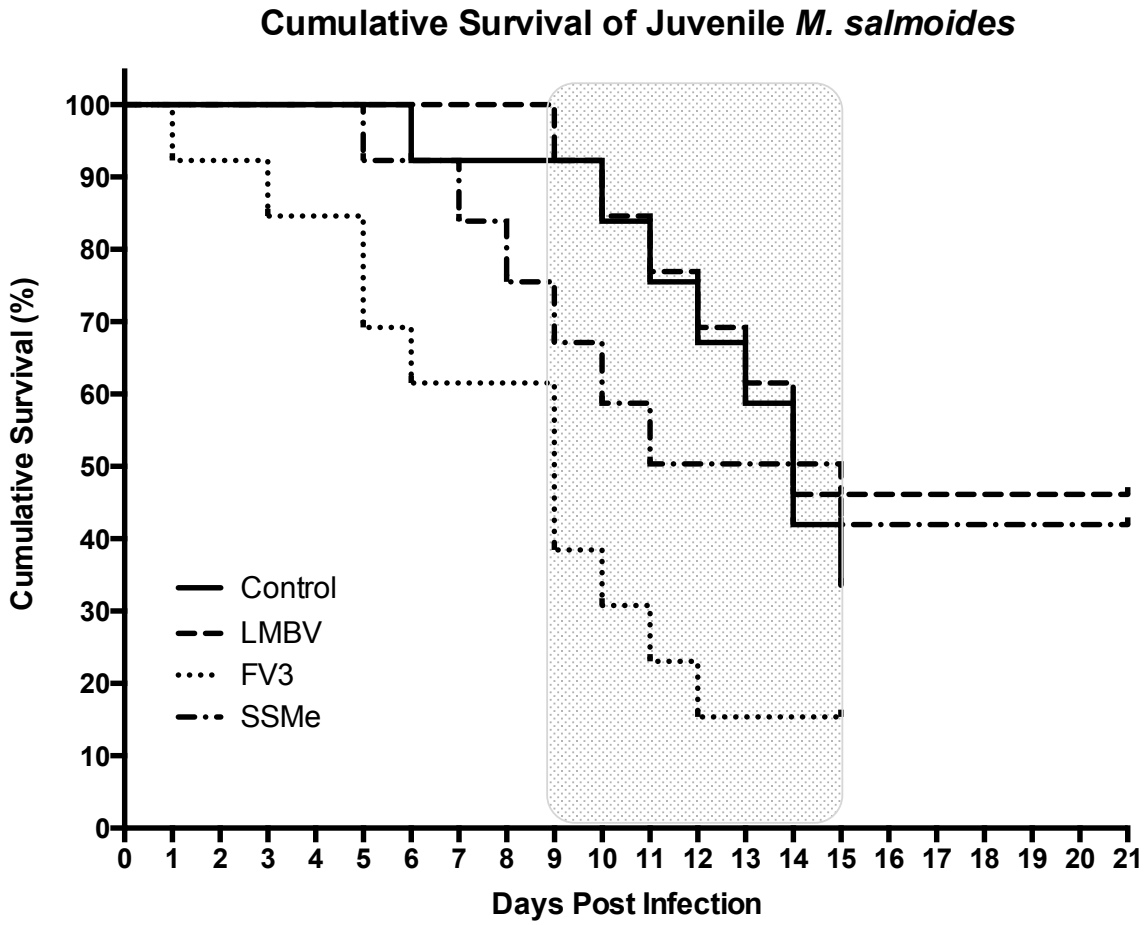

Figure 1.2 Challenge I: Percent cumulative survival of *M. salmoides* juveniles infected with FV3, SSMe, and LMBV. Survival for the negative controls is shown as a solid line. The gray box highlights the period when visible signs of a fungal infection were present in all tanks. During this time, fish received daily doses of the antiparasitic MGF treatment. Significant differences in the survival curves were calculated using a Mantel-Cox log-rank test. No significant difference was found between control and LMBV treatments ($p = 0.6259$), control and SSMe treatments ($p = 0.9271$), FV3 and SSMe treatments ($p = 0.0917$), and LMBV and SSMe treatments ($p = 0.5516$). Significant differences were found between control and FV3 treatments ($p = 0.0233$) and FV3 and LMBV treatments ($p = 0.0077$).

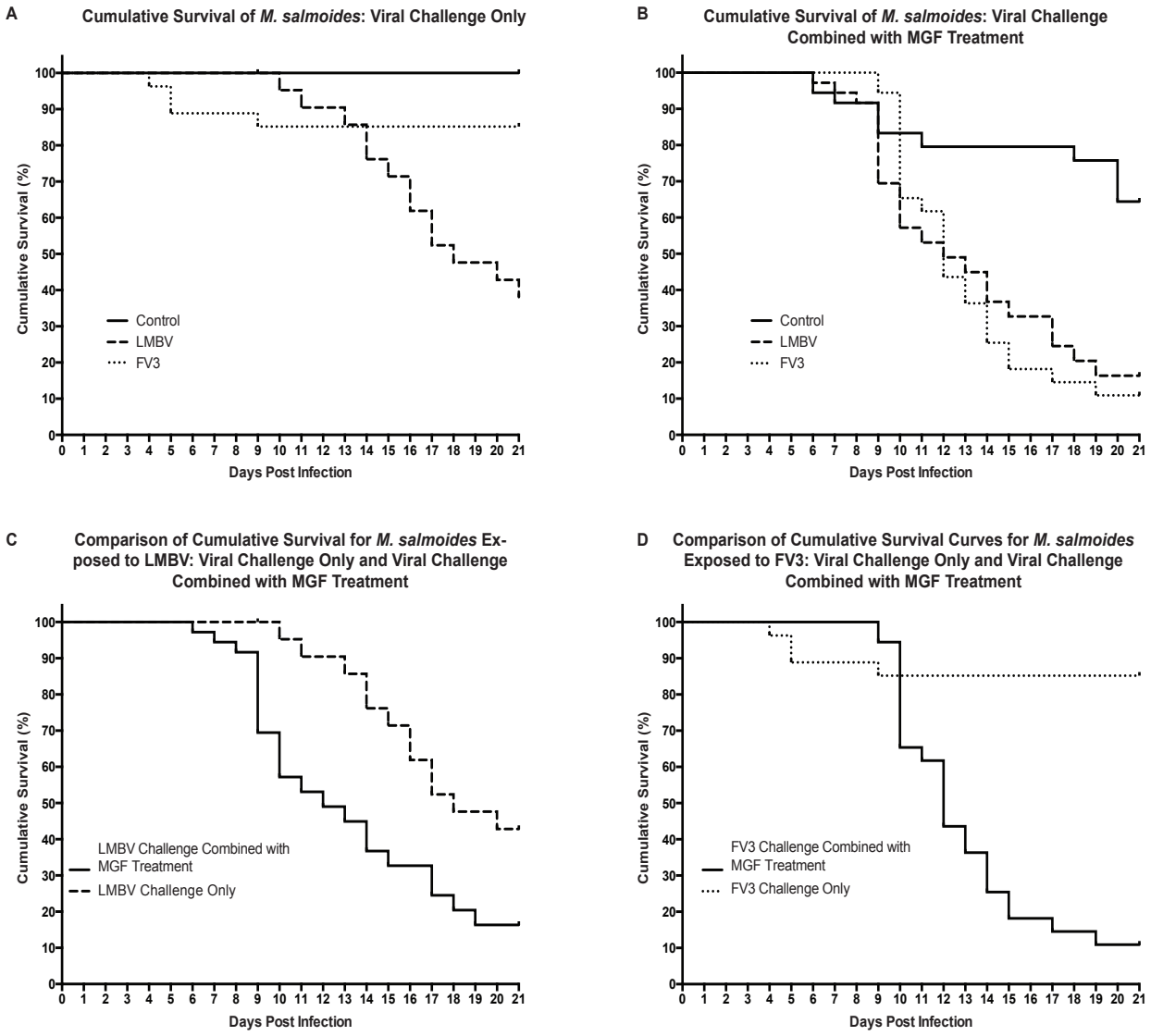

Figure 1.3: Challenge II: Percent cumulative survival of *M. salmoides* juveniles exposed to virus only and those exposed to virus in combination with MGF treatment. See Table 1.3 for survival curve comparison results based on Mantel-Cox log-rank tests. Panel A) Cumulative survival for fish infected with LMBV and FV3 in the absence of antiparasitic MGF treatment. Cumulative survival for the negative controls is shown as a solid line. Panel B) Cumulative survival for fish infected with LMBV and FV3 in the presence of daily antiparasitic MGF treatment. Cumulative survival for the negative controls is shown as a solid line. Panels C and D) Cumulative survival for LMBV and FV3 treatments under both virus challenge only and virus challenge combined with MGF treatment, respectively. Significant differences were detected between LMBV treatment curves (panel C, $p = 0.0031$) and between FV3 treatment curves (panel D, $p < 0.0001$).

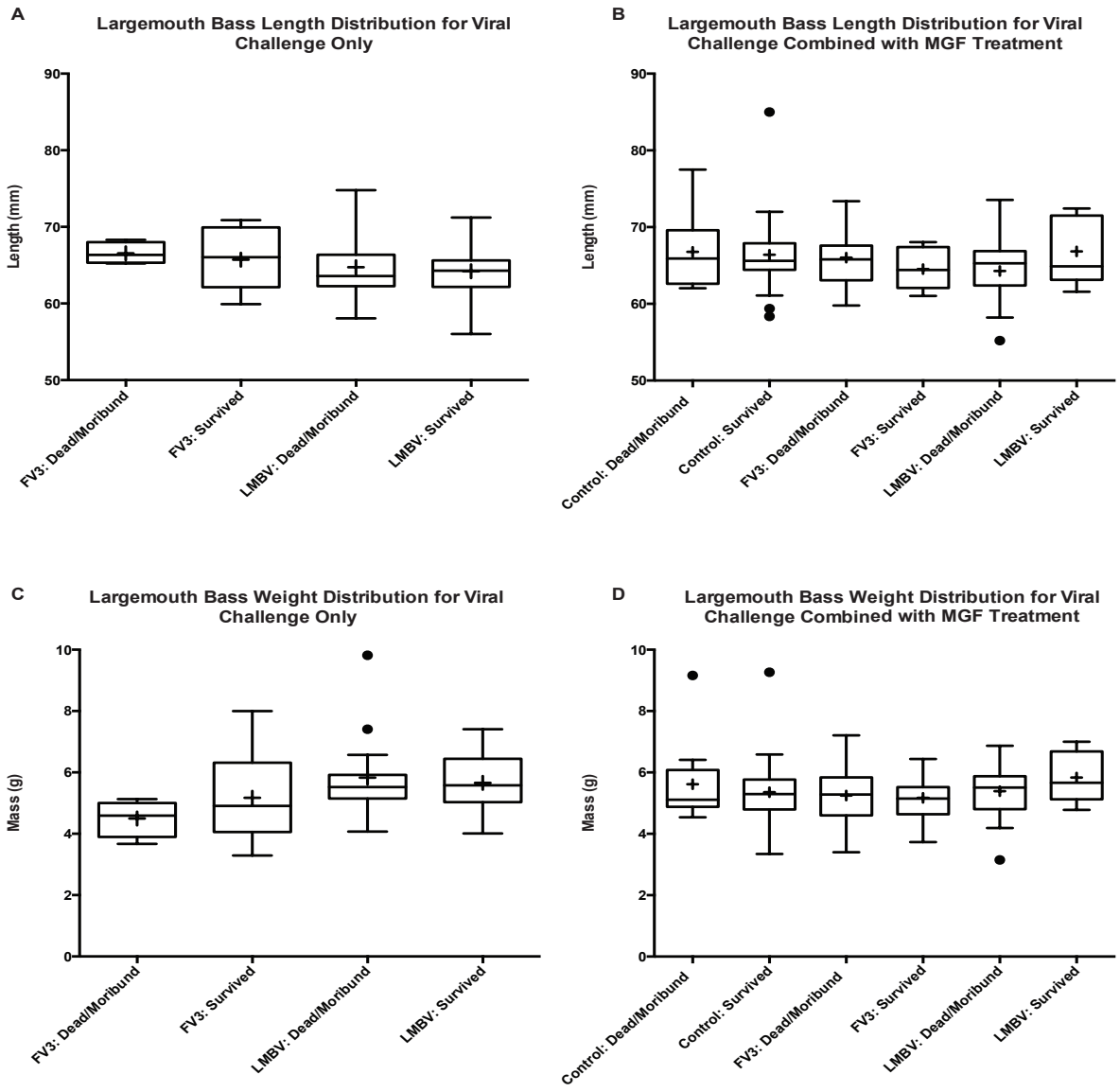

Figure 1.4: Length and weight distributions for dead/moribund and surviving *M. salmoides* juveniles exposed to virus only and those exposed to virus in combination with MGF treatment. Panels A and C) Standard length and weight distributions for dead/moribund and surviving fish infected with virus only. All fish in the control treatment survived until the end of the experiment, so the data was excluded from the analysis. Kruskal-Wallis (medians) and ANOVA tests (means) were performed, and no significant differences in the length and mass values were detected between dead/moribund fish and those that survived. Panels B and D) Length and weight distributions for dead/moribund and surviving fish exposed to virus in combination with MGF treatment. Kruskal-Wallis (medians) and ANOVA tests (means) were performed, and no significant differences in the length and mass values were detected between dead/moribund fish and those that survived.

Chapter 2: Recent host-shifts in ranaviruses: signatures of positive selection in the viral genome

ABSTRACT

Ranaviruses have been implicated in recent declines in global amphibian populations. Compared to the larger *Iridoviridae* family to which they belong, ranaviruses have a wide host range in that species/strains are known to infect fish, amphibians, and reptiles, presumably due to recent host switching events. We used eight sequenced ranavirus genomes and two selection-detection methods (site-based and branch-based) to identify genes that exhibited signatures of positive selection, potentially due to the selective pressures at play during host switching. We found evidence of positive selection acting on four genes via the site-based method, three of which are newly-acquired genes unique to ranavirus genomes. Using the branch-based method, we identified eight additional candidate genes that exhibited signatures of *dN*/*dS* greater than one in the clade where intense host switching has occurred. We find that these branch-specific patterns of elevated *dN*/*dS* are enriched in a small group of viral genes that have been most recently acquired in the ranavirus genome, compared to core genes that are shared among all *Iridoviridae*. Our results suggest that the group of newly acquired genes in the ranavirus genome may have undergone recent adaptive changes that have facilitated interspecies and interclass host switching.

INTRODUCTION

Since the 1980s, amphibian populations have increasingly declined on a global scale, and by 2004, at least 32% of all amphibian species were categorized as threatened

by the International Union for Conservation of Nature (Collins & Storfer 2003; Stuart *et al*. 2004). In addition to several biotic and abiotic factors, two emerging pathogens, a chytrid fungus (*Batrachochytrium dendrobatidis*) and a virus (ranavirus), have been implicated in this ecological crisis (Daszak *et al*. 1999). *Ranavirus* is a member of the *Iridoviridae* family, which consists of large dsDNA viruses that are currently divided into five genera (Williams *et al*. 2005). Two of these infect invertebrates (*Iridovirus* and *Chloriridovirus*), and three infect heterothermic vertebrates (*Lymphocystivirus, Megalocytivirus*, and *Ranavirus*). *Lymphocystivirus* and *Megalocytivirus* solely infect fish, but ranaviruses infect fish, amphibians, and reptiles, presumably due to recent host switching events (Jancovich *et al*. 2010). Direct contact and ingestion of contaminated water are known transmission modes, and infection leads to necrosis of the liver, spleen, skin, and haematopoietic tissue (Lesbarreres *et al*. 2012). The bait, pet, and food industries are major factors in the spread of these pathogens geographically to naïve localities and species (Jancovich *et al*. 2005; Picco & Collins 2008). Although both pathogens pose a lethal threat to amphibians, ranaviruses tend to cause recurrent population die-offs that target larval stage amphibians (Chinchar 2002). Numerous ranavirus species/strains have contributed significantly to amphibian die-offs on five continents, and are known to infect over 70 amphibian species (Daszak *et al*. 1999; Miller *et al*. 2011).

Ranaviruses are emerging pathogens, meaning amphibian populations have only recently been affected due to increased incidence, geography, or host range (Rachowicz *et al*. 2005). There is significant evidence for host switching of ranaviruses between vertebrate classes (Jancovich *et al*. 2005, 2010; Bandin & Dopazo 2011) as well as documented instances of simultaneous infections of sympatric fish and frogs with the same viral isolate (Mao *et al*. 1999; Bayley *et al*. 2013). In viral host switching, the virus

of one species evolves such that it can replicate and spread in a new species. There is great interest in understanding this process, not just for conservation reasons, but also because it underlies the emergence of new human diseases (Wain *et al*. 2007; Parrish *et al*. 2008; Choe *et al*. 2011; Demogines *et al*. 2012a, 2012b, 2013). In general, host switching is thought to occur via a series of steps that involve both geographic movement of viruses (leading to exposure to new hosts), and in most cases, the evolution of viral genomes for compatibility with new host species (Herfst *et al*. 2012; Kaelber *et al*. 2012). With regard to host switching in ranaviruses, the process probably began with the exposure of a novel amphibian host to a viral strain that originated in a fish host (Figure 2.1). New mutations, including point mutations, insertions, deletions, and even the acquisition or loss of whole genes, may have made a certain strain of the fish virus compatible with the new amphibian host species. Following initial infection, additional mutations or gene acquisitions may have increased viral fitness in a new host, thus allowing local transmission within species and between classes. Indeed, ranavirus genomes have diverged in sequence, and are known to have acquired several new genes that are not found in the broader *Iridoviridae* family (Eaton *et al*. 2007). These new genes may have been acquired via the processes of horizontal gene transfer or recombination. The presence of several virus-encoded cellular homologs in ranavirus genomes supports the acquisition of new viral genes by horizontal transfer from host genomes (Tidona & Darai 2000). In addition, the presence of several ranavirus genes without corresponding cellular homologs, coupled with the high rate of recombination seen in these viruses, suggests that recombination is also a significant source of new genes (Eaton *et al*. 2007, Chinchar *et al*. 2011). The situation with ranaviruses may be analogous to the evolution of simian immunodeficiency viruses, where the continued

acquisition and evolution of new viral genes led to expansion of these viruses to new primate hosts and ultimately to humans, giving rise to HIV (Gifford 2011).

The main factor limiting our ability to understand the broad host range of ranaviruses is the paucity of information regarding the viral genes that govern host-specificity. Ranavirus genomes encode approximately 100-140 putative gene products, most of which have unknown functions (Grayfer *et al*. 2012). Computational tools that can detect episodes of rapid adaptation of specific genes, in conjunction with the increased availability of genomic data, have provided new avenues for identifying candidate genes potentially involved in dynamic processes such as host switching (Sawyer & Elde 2012). In many cases, genes that are involved in host-pathogen interactions exhibit episodes of intense positive selection during the process of establishing a new host species (Shackleton *et al*. 2005; Hoelzer *et al*. 2008; Meyerson & Sawyer 2011; Bhatt *et al*. 2013). Genes identified via these methods are often involved in immune evasion, replication, reproduction, gene expression, host-pathogen coevolution, and host defenses (Endo *et al*. 1996; Yang *et al*. 2000; McLysaght *et al*. 2003; Harrison & Bonning 2004; Shackleton *et al*. 2005; Sabeti *et al*. 2006; Kosiol *et al*. 2008; Elde *et al*. 2012). However, the absence of positive selection does not necessarily mean that genes are not key factors in host range and immune evasion, just as the presence of positive selection does not automatically assign genes to these roles.

Methods used to study adaptive molecular evolution typically involve the identification of gene regions or specific codons where the nonsynonymous substitution rate (*dN*) exceeds the synonymous substitution rate (*dS*). This is expected only when nonsynonymous mutations offer a selective advantage and become fixed at a higher rate than expected under neutral theory (Goldman & Yang 1994). In practice, $dN/dS > 1$, $dN/dS = 1$, $dN/dS < 1$ are inferred to reflect positive selection, neutral evolution, and

purifying selection respectively. Two widely used selection-detection methods identify individual codon sites (site-based model) and specific branches along a phylogeny (branch-based model) that exhibit signatures of positive selection (Yang 1998; Nielsen & Yang 1998). The availability of several complete *Iridoviridae* genomes from viruses infecting vertebrates provides the opportunity to assess the prevalence of positive selection within specific genes and across the ranavirus phylogeny.

Here, we describe a genomic-scale characterization of ranavirus evolution. The goals of this study were to detect viral genes that are under positive selection, particularly on branches in the phylogeny where potential host switching events have occurred. Using eight fully sequenced ranavirus genomes, we analyzed 46 genes for evolutionary signatures. We found evidence of positive selection acting on four genes via the site-based method, three of which are newly acquired genes unique to ranavirus. Using the branch-based method, we identified eight additional candidate genes that exhibited signatures of $dN/dS > 1$ along at least one branch in the clade where intense host switching has occurred. We also found evidence of significantly higher dN/dS values for ranavirus-specific genes along individual branches when compared to core genes shared among all *Iridoviridae*. We conclude that adaptive evolution is occurring in newly acquired viral genes, and propose that both the acquisition and the subsequent evolution of these genes has been key to facilitating host switching of these viruses.

METHODS

Data Collection and Recombination Analyses

Sequences for 75 open reading frames (ORFs) from the fifteen complete vertebrate Iridoviridae genomes were downloaded from Genbank (Table 2.1). These

ORFs were placed into specific gene groups based on the genomic analysis of Eaton *et al*. (2007), which identified orthologous genes throughout the family using a combination of BLAST searches and Viral Orthologous Clusters software (Eaton *et al*. 2007). Genes from viral genomes not analyzed in Eaton *et al*. (2007) were identified via BLAST searches and analyses conducted by the authors of the published genome sequences (Huang *et al*. 2009; Jancovich *et al*. 2010; Lei *et al*. 2012; Mavian *et al*. 2012a; Mavian *et al*. 2012b). The first group, the Core gene group, included 26 ORFs that have orthologs in all members of the *Iridovirdae* family (including the two invertebrate genera *Iridovirus* and *Chloridovirus*). The second group, the RV gene group, included 23 ORFs that have orthologs only found in *Ranavirus*. The third group, the F/A/R gene group, included 6 ORFs that have orthologs only found in the Target Clade that includes the FV3, STIV, RGV, TFV, CMTV, ATV, ESV, and EHNV species/strains. The remaining ORFs were identified as "Other" and included 20 ORFs that are shared between ranavirus species/strains and at least one other genus (including those that infect invertebrates).

The sequences for each ORF were aligned using Clustal W (Larkin *et al*. 2007). Four ORFs with alignments that were judged to be unreliable were excluded from further analyses (Supplementary Table S1). CodonTest (Delport *et al*. 2010) was then used to determine the best nucleotide substitution model for the remaining 71 alignments. The nucleotide substitution models were then used to calibrate the program GARD (Genetic Algorithm for Recombination Detection) (Pond *et al*. 2006) that was used to test for the presence of recombination in each ORF. A total of 25 ORFs with significant evidence of recombination were split based on the location of the recombination breakpoints, and each component was individually tested in subsequent analyses.

Phylogenetic Reconstruction

Phylogenetic trees were reconstructed for individual ORF alignments using the rapid bootstrap command in RAxML, which automatically applies the GTR + Γ model (500 bootstrap replicates were used to assess node support). Preliminary analyses of alignments that included sequences from the fifteen *Iridoviridae* and ten ranavirus species/strains resulted in highly divergent phylogenies that fell outside of the ideal range for the selection analyses (with some divergences greater than 20 nucleotide substitutions per codon). To analyze divergences closer to one nucleotide substitution per codon, which is a commonly used standard (Anisomova 2001), only alignments using the eight sequences from the Target Clade were used in subsequent analyses. Please note that the apparent topological differences in the 15-taxon tree and the 8-taxon trees are due to the inclusion of additional data and the placement of the root. Supplemental Figure S2.4 illustrates that the relationship between (ATV, (EHNV, ESV)) is consistent between the unrooted topologies. Moreover, the rooted topology seen in Supplemental Figure S2.4b can be recovered in Supplemental Figure S2.4a when the root of the 15-taxon tree is moved to match that of the 8-taxon tree. Thus, the phylogenetic relationships are consistent among the phylogenies.

The resulting eight-taxon phylogenies were accepted based on the criteria of a fully resolved topology and bootstrap support values that were greater than or equal to 50 for the following clades: (FV3, STIV, RGV), (FV3,STIV,RGV,TFV,CMTV), and (ATV, EHNV, ESV). A total of 25 ORFs that did not meet these requirements were excluded from the analyses, and the remaining Core (18), RV (11), F/A/R (6), and Other (11) ORFs were further analyzed. The resulting topologies for each ORF, without the inclusion of branch lengths, were then used as input topologies for the selection detection

analyses in the PAML4 (Phylogenetic Analysis by Maximum Likelihood) software package (Yang 2007).

To determine the overall phylogenetic relationships of the vertebrate *Iridoviridae* viruses, a phylogenetic tree was reconstructed via maximum likelihood using a concatenated dataset of the 26 Core group genes (Figure 2.1; Supplemental Figure S2.1). The phylogeny was reconstructed using RAxML (Stamatakis 2006) under the GTR + Γ model with the data partitioned by gene, and 500 bootstrap replicates were used to assess node support.

Detection of Positive Selection

Identifying Selection at Specific Sites

Site-based models were used to identify specific sites (codons) under positive selection via the NSsites models implemented in PAML4. The NSsites models allow dN/dS to vary among sites, and they use likelihood ratio tests (LRT) and Bayesian analyses to identify positively-selected sites. Each test involves the comparison of a null model (M8a) to a model that incorporates positive selection (M8). M8a versus M8 models were run for each ORF, and the significance of the likelihood-ratio scores was determined via the use of a chi-square distribution ($\alpha = 0.05$). The models were run under the f61 codon frequency models. Sites that fell into positively-selected classes were then subjected to a Bayes empirical Bayes (BEB) analysis, which calculates a posterior probability of a site fitting into a $dN/dS>1$ class. The sites detected were reported as under positive selection if they had a posterior probability of greater than 0.80 under the BEB method.

Identifying Selection along Branches

Branch-based models seek to identify specific branches along a phylogeny that exhibit signatures of positive selection. To test for positive selection along specific branches in the phylogeny, the free-ratio model that is implemented in PAML4 was used to estimate branch-specific *dN*/*dS* ratios. The free-ratio model allows a unique *dN*/*dS* value to be calculated for each branch in the phylogeny. The free-ratio model was run for each ORF under the f61 codon frequency model to account for codon usage bias. To maximize topological consistency in our analyses, we constrained the topology of gene trees with alternate topologies to the topology supported by Figure 2.2 (accepted topology). We then ran the branch-based analysis and used likelihood ratio tests to determine if the results were significantly different between the two trees. The accepted topology is presented in cases where no significant difference was detected. Likelihood-ratio tests were performed to determine if the likelihood value for each branch that was assigned *dN*/*dS* >1 was statistically significant when compared to the likelihood value of the same branch when fixed at *dN*/*dS*=1. The *dN*/*dS* of the target branch was fixed at 1, and the likelihood was calculated (null model). The null model was then compared to the alternative model, which fixed the target branch at the *dN*/*dS* value calculated under the free-ratio model.

To determine if there were significant differences in *dN*/*dS* values among the Core, RV, and F/A/R gene groups, we compared both the median and mean *dN*/*dS* values for each group for each branch using Kruskal-Wallis and ANOVA tests, respectively. The *dN*/*dS* values of several branches were undefined due to zero observed synonymous changes along the branch (*dN*/0). However, branches with $S = 0$ are not uninformative with respect to selection; they are likely to be evolving near neutral if N is small, and under positive selection if N is large. Therefore, to include these branches in our analyses,

we assigned a value of *dN*/*dS*= 0.95 (near neutral) for branches with observed *N*:*S* ratios greater than 1:0 and less than 4:0. Likewise, we assigned a value of *dN*/*dS*= 2.0 to branches with observed *N*:*S* values greater than 4:0 to account for likely positive selection. We also obtain consistent results where a *dN*/*dS* value of 1.5 is assigned, instead of 2.0, to this latter class. These assignments were consistent among all gene groups to keep the relative differences of the median and mean *dN*/*dS* values consistent for each data set.

Protein Structure Prediction Analyses

The secondary protein structure for each of the ORFs found to be under positive selection under the NSsites method was determined using the Phyre2 server (Kelley and Sternberg 2009). The full-length FV3 amino acid sequence for each ORF was used as the input sequence for the analysis. The predicted domains, secondary structure, and disorder were determined and a relative confidence value for each component was calculated. The positively selected sites identified by the NSsites methods were then mapped onto the predicted secondary structure of the protein.

RESULTS

In previous studies the full genome sequences of 15 vertebrate *Iridoviridae* viruses have been determined (Table 2.1). Ten of these are ranaviruses, and represent viruses isolated from bony fish, salamander, frog, and turtle species (Figure 2.2, Supplemental Figure S2.1). These viral genomes vary somewhat in length and gene content, and contain approximately 100-140 genes. The genes examined in this analysis were identified previously as orthologs via a whole-genome comparative analysis (Eaton

et al. 2007; Methods). In total, we were able to align 75 orthologs from these genomes, and assigned each of these 75 genes to four groups depending on when each gene was acquired in the phylogenetic history of these viruses (Figure 2.2). The "Core" gene group includes 26 genes found in all 15 sequenced Iridoviridae genomes. The "RV" (ranavirus) gene group comprises 23 genes specific to the 10 ranavirus genomes, but not found in the other 5 genomes. This group of genes was gained prior to the initial host switch from fish to amphibians (Jancovich *et al*. 2010). The "F/A/R" (fish/amphibian/reptile) gene group includes 6 genes found exclusively in the Target Clade, which is the monophyletic group of ranaviruses, excluding Singapore Grouper Iridovirus (SGIV) and Grouper Iridovirus (GIV). The Target clade includes the group of viruses that has been newly transmitted beyond fish. This F/A/R gene group is particularly interesting, because these genes are specific to the clade of viruses that has experienced the most recent host switching. Lastly, the remaining 20 genes were assigned to the "Other" gene group, which included genes found in ranaviruses and at least one other *Iridoviridae* genus (i.e., *Ranavirus* and *Chloriridovirus* but no other genera).

Four Genes Under Positive Selection Identified Using Site-based Tests

We next analyzed a subset (46/75 genes) of the individual gene alignments for patterns of non-synonymous and synonymous substitution using the NSsites codon models in PAML (Yang 1997). The 29 genes excluded from the analysis failed to meet the outlined criteria based on sequence alignment and phylogenetic reconstruction (Methods; Supplemental Table S2.1). Of the 46 analyzed, we identified four genes that exhibited signatures of positive selection under the NSsites models (Table 2.2). This group of genes includes a single member of the Core gene group (ORF53R), one gene from the RV gene group (ORF79R), and two genes from the F/A/R gene group (ORF40R

and ORF71R). Evidence of selection acting on RV and F/A/R genes might be explained by the fact that these genes are recently acquired and unique to a single genus where host switching is evident. Relative to the newly acquired RV and F/A/R genes, evidence of selection acting on Core genes is unexpected due to the highly conserved nature of this group of genes within *Iridoviridae*. The conservation of Core genes may indicate that, in general, these genes are essential to the viral life cycle and that they are subject to purifying selection pressures.

In addition to identifying individual genes that have experienced positive selection, specific codons under positive selection were also identified (Table 2.2). At least one codon with a posterior probability of >0.80 was identified for each gene. To determine whether the positively-selected sites fell into structurally flexible regions of the protein, we utilized the Phyre2 server to predict the secondary structure of the proteins encoded by the four genes that exhibited signatures of positive selection under the NSsites model. We then mapped the amino acid sites onto the secondary structure to determine if any of the residues resided in areas of conformational flexibility, described as unstructured regions with high disorder (Supplemental Figure S2.2). We found that 1/1 site for ORF53R, 10/10 sites for ORF79R, 3/8 sites ORF40R, and 1/4 sites for ORF71R fell into unstructured (neither α-helix nor β-strand) regions of the protein. Unstructured regions of proteins may be more tolerant to conformational and adaptive changes (Nilsson *et al*. 2011).

Eleven Genes with *dN/dS* >1 Identified Using Branch-Based Tests

Based on the evolutionary relationships of the viral strains and the species from which they were isolated, we can support hypotheses about where host switching events occurred in the ranavirus phylogeny (Jancovich *et al*. 2010). For instance, we can support

the hypothesis that the original host switch involved a jump from fish to salamanders (Figure 2.2, point A). Moreover, we can also support a host switch from fish to frogs or from fish to frogs and turtles on the lineage leading to the (FV3,STIV,RGV,TFV,CMTV) clade (Figure 2, point B). To examine evidence for accelerated gene evolution of any genes on these lineages, we next analyzed branch-specific patterns of *dN/dS* for each of our 46 genes. Whereas the previous test for selection identified specific codons that have *dN/dS* >1, in this analysis each gene was assigned a global *dN/dS* value along every branch of the tree.

We identified 11 genes that exhibited signatures of *dN/dS* >1 along at least one branch under the free-ratio test. These included 3 genes from the RV and Core gene groups (Figure 2.3), 4 genes from the F/A/R gene group (Figure 2.4), and 4 genes from the Other gene group (Supplemental Figure S2.3). In some cases, *dN/dS* values were >> 1, as for gene ORF40R on the branch leading to Tiger Frog Virus (TFV), where *dN/dS* = 3.7 (Figure 2.4b). In other cases, *dN/dS* values were closer to one, and may reflect neutral evolution rather than positive selection. Likelihood-ratio tests were performed to determine if the likelihood values obtained with individual branches assigned *dN/dS* >1 are statistically greater than likelihood values obtained when the same branch is fixed at *dN/dS*=1. Based on this analysis, on no branch is *dN/dS* significantly greater than one; however, the identification of genes with elevated *dN/dS* signatures is notable considering *dN/dS* is averaged across the entire gene. There is reason to believe that some of these signatures could still be meaningful, because three of these same genes (ORF79R, ORF40R, and ORF71R) were also previously identified as evolving under positive selection in the site-based analysis described above (Table 2).

Despite the fact that F/A/R represents the smallest gene group analyzed, four of the six F/A/R genes analyzed (ORF39R, ORF40R, ORF71R, and ORF93L) exhibited

dN/*dS* >1 on at least one branch (Figure 2.4a-d). When compared to the percentage of genes with *dN*/*dS* >1 branches for the Core (5%), RV (18%), and "Other" (36%) groups, genes from the F/A/R group (67%) exhibit a relatively higher percentage of genes with elevated *dN*/*dS* values. Based on this result and the identification of two out of the six F/A/R genes via the site-based method (Table 2.2), we conducted an enrichment analysis to determine the probability that four randomly selected genes would belong to the F/A/R gene group. We calculated probabilities of p = 0.00009 when 46 genes were included and p = 0.00001 when 75 genes were included in the analysis. These results are consistent with selection in ranaviruses having been biased towards viral genes that are new to this clade.

Significant Differences in Median and Mean *dN*/*dS* Values of Core, RV, and F/A/R Gene Groups Detected Using Branch-Based Tests

We wished to further examine whether there are significant differences in branch *dN*/*dS* values among the Core, RV, and F/A/R gene groups. For each of the 13 branches in the Target Clade (the clade where host switching is pervasive; Figure 2.2), median and mean *dN*/*dS* values for each gene group were compared. For each branch, we performed a Kruskal-Wallis (K-W) test ($\alpha = 0.05$) to compare the median *dN*/*dS* values for each gene group, and a one-way ANOVA ($\alpha=0.05$) to compare the mean values. On 4 of the 13 branches, significant differences were noted (Figure 2.5). We detected significant differences in the medians and means for Branch B (K-W p = 0.001 and ANOVA: p = 0.0001) and the ATV branch (K-W: p = 0.0037 and ANOVA: p = 0.0005) (Figure 2.5b, d). In both cases, the median and mean *dN*/*dS* values were highest for the F/A/R gene group. However, while F/A/R *dN*/*dS* values were significantly different from the Core gene group values, they did not differ significantly from the RV gene group based on

post-hoc tests. We also detected a significant difference in the medians for the TFV branch ($p = 0.0381$) using the K-W test (Figure 2.5c); however, the ANOVA was not significant ($p > 0.05$). Lastly, Branch A was marginally significantly different in the ANOVA analysis ($p = 0.0446$), but there was no significant difference in the medians based on the K-W test ($p > 0.05$). Higher *dN*/*dS* values for the ranavirus-specific gene groups along Branch A and Branch B could indicate selection acting prior to a potential host switch between amphibians and reptiles, which is the most recent putative host switching event examined in this study. Moreover, this interpretation might also apply to the ATV branch, which corresponds to selection following a host switch from fish to amphibians. The trend of higher mean and median values for the F/A/R gene group for three of the branches (Figure 2.5b-d) is consistent with the idea that F/A/R genes are involved in ranavirus host switching.

DISCUSSION

The phylogeny of *Iridoviridae* viruses, particularly in the ranavirus clade, is consistent with extensive host switching by these viruses (Figure 2.2). The adaptation of a virus to a novel host species is often paired with intense episodes of positive selection acting on genes that increase the compatibility between the virus and host. We conducted a genome-wide scan for genes and lineages under positive selection to gain further insights into ranavirus host switching. We identified a total of twelve genes exhibiting signatures of *dN*/*dS* >1, and six of these are in a special category of genes, which have been newly acquired by these viruses (RV and F/A/R gene groups). We expected that genes in the RV and F/A/R gene groups may play a fundamental role in the adaptation of ranaviruses to new hosts, because the acquisition of these genes appears to have been

coincident with the onset of host switching in this virus family (Jancovich *et al*. 2010). In support of this hypothesis, several approaches show that dN/dS tends to be elevated in these gene families.

Regarding the six ranavirus-specific genes that we identified as adaptively evolving, several aspects are notable. ORF79R (RV gene group) was the only gene to have dN/dS>1 on the branch that separates the two major ranavirus clades (Figure 2.3c,), which suggests gene evolution as the viruses spread to new frog species (Figure 2.2). Thus, ORF79R may represent a gene that was fundamental to the initial host switch from fish to salamanders and possibly subsequent jumps. The identification of four genes specific to the F/A/R gene group with *dN/dS*>1 on at least one branch supports the hypotheses of intensified positive selection acting on these newer genes. Three of the four identified F/A/R genes are of unknown function and have no known homologs. It is possible that these genes encode novel virulence factors, which are typically proteins that counteract host-specific immune defenses. For instance, examinations of *Poxviridae* and *Herpesvirales* genomes have revealed the presence of virus-encoded homologs of cellular cytokines that have evolved to modulate the host immune response (Slobedman *et al*. 2009). In fact, the F/A/R gene ORF26R (alpha subunit of eukaryotic initiation factor 2) is a virus-encoded cellular homolog (Eaton *et al*. 2007). Although no evidence of positive selection was detected in ORF26R, a recent study using a knockout approach showed that the impairment of the gene in FV3 resulted in decreased virulence *in vivo* (Chen *et al*. 2011). We also found that mean and median *dN/dS* values for several branches in the ranavirus phylogeny were significantly higher for the F/A/R group when compared to the Core gene group (Figure 2.5). Considering this result and the low probability ($p = 0.00009$) of finding *dN/dS* >1 branches in four of the six F/A/R genes by random chance, we can conclude that F/A/R genes represent a group of newly acquired viral genes that

have undergone adaptive changes that may have facilitated interspecies and interclass host switching (Figure 2.1).

There are other examples where adaptive evolution of virus genomes has accompanied host switching events. For instance, parvoviruses experienced adaptive evolution of the gene encoding the surface-exposed viral protein, capsid, after this virus first became established in dogs in the 1970s (Hoelzer *et al*. 2008). These evolutionary changes improved interaction between capsid and the dog-encoded cell surface receptor for virus entry, transferrin receptor. Also, a recent study of the adaptive evolution of influenza viruses showed that elevated adaptive evolution was associated with the establishment of avian influenza viruses in swine populations (Bhatt *et al*. 2013). In this case, signatures of selection were located across the genome and so the specific host-virus interactions that drove this evolution are unknown. With regard to ranaviruses, this situation is unique from these other examples because the cross-species transmission of these viruses appears to have coincided with both the addition of new genes to the viral genome, as well as their evolutionary refinement for infection of new host species.

Although we did not explicitly test the hypothetical model presented in Figure 2.1, the relatively recent acquisition of new genes (RV and F/A/R genes) coupled with an excess of nonsynonymous mutations in these genes supports the idea that gene gain followed by gene-specific mutation may influence ranavirus host switching. Moreover, evidence of simultaneous infections of sympatric fish and amphibian species with the same viral isolates (Mao *et al*. 1999) supports the idea that both interspecies and interclass transmission may be a common occurrence. However, additional studies of ranavirus spread in natural environments are needed to specifically test the model. Furthermore, estimates of viral divergence times and the timing of gene acquisitions are needed to fully understand the emergent nature of ranaviruses. Although humans have

influenced the movement of aquatic species for millennia, it is more likely that recent human activities (globalized food, bait, and pet trades) have enhanced the spread of these pathogens to naïve species. Unfortunately, since the rates of evolution in *Iridoviridae* are currently unknown, uses of molecular clock methods are not possible (Ridenhour & Storfer 2008). In the absence of this data, the genes identified in this study are prime candidates for knockout and site-directed mutagenesis analyses to determine gene function and the effect of the genes on viral host-range in this unique genus.

Viral Species/Strain	Genus	Host Species	Genbank Accession No.
Frog virus 3 (FV3)	*Ranavirus*	*Rana pipiens*	AY548484 (Tan *et al*. 2004)
Soft-shelled turtle iridovirus (STIV)	*Ranavirus*	*Trionyx sinensis*	EU627010 (Huang *et al*. 2009)
Rana grylio virus (RGV)	*Ranavirus*	*Rana grylio*	JQ654586 (Lei *et al*. 2012)
Tiger frog virus (TFV)	*Ranavirus*	*Rana tigrina*	AF389451 (He *et al*. 2001)
Common midwife toad virus (CMTV)	*Ranavirus*	*Alytes obstetricans*	JQ231222 (Mavian *et al*. 2012a)
Ambystoma tigrinum virus (ATV)	*Ranavirus*	*Ambystoma tigrinum*	AY150217 (Jancovich *et al*. 2003)
Epizootic haematopoietic necrosis virus (EHNV)	*Ranavirus*	*Oncorhynchus mykiss*	FJ433873 (Jancovich *et al*. 2010)
European sheatfish virus (ESV)	*Ranavirus*	*Silurus glanis*	JQ724856 (Mavian *et al*. 2012b)
Singapore grouper iridovirus (SGIV)	*Ranavirus*	*Epinephelus tauvina*	AY521625 (Song *et al*. 2004)
Grouper iridovirus (GIV)	*Ranavirus*	*Epinephelus awoara*	AY666015 (Tsai *et al*. 2005)
Lymphocystis disease virus 1 (LCDV1)	*Lymphocystivirus*	*Platichthys flesus*	L63545 (Tidona and Darai 1997)
Lymphocystis disease virus China (LCDVC)	*Lymphocystivirus*	*Paralichthys olivaceus*	AY380826 (Zhang *et al*. 2004)
Infectious spleen and kidney necrosis virus (ISKNV)	*Megalocytivirus*	*Siniperca chuatsi*	AF371960 (He *et al*. 2002)
Orange-spotted grouper virus (OSGIV)	*Megalocytivirus*	*Epinephelus coioides*	AY894343 (Lu *et al*. 2005)
Rock bream iridovirus (RBIV)	*Megalocytivirus*	*Oplegnathus fasciatus*	AY532606 (Do *et al*. 2004)

Table 2.1: Fifteen genomes utilized in the present study and their host species.

Gene Name	Gene Group	M8a-M8 2Δl	p-value (α=0.05)	Tree Length	*,§AA Positions of dN/dS>1 *PP>0.95, **PP>0.99
ORF53R/ Myristylate membrane protein	Core	3.92	0.0478	0.16	387G (0.86)
ORF79R/ Putative ATPase-dependent protease	RV	5.19	0.0227	1.06	81L (0.88), 82A*, 84Q (0.89), 86L (0.81), 87V (0.83), 91S (0.88), 123L (0.85), 137R*, 140N (0.87), 146L (0.83)
ORF40R	F/A/R	4.74	0.0294	0.82	29M (0.84), 131Q*, 134A (0.88), 164T (0.83), 165V (0.81), 168L (0.83), 191T (0.88), 192P (0.86)
ORF71R	F/A/R	12.9	0.0003	0.38	19T**, 50M (0.94), 58I**, 70I*

Table 2.2: Four genes under positive selection identified using site-based tests.

Gene Group	Gene Name	FV3 Designation	Status	Recombination
Core Gene Group	ATPase-like protein	15R	Excluded based on phylogeny	Yes
Core Gene Group	D5 family NTPase involved in DNA replication	22R	Excluded based on phylogeny	Yes
Core Gene Group	Deoxynucleoside kinase	85R	Used in analysis	N/A
Core Gene Group	DNA -dep RNA pol-II largest subunit	8R	Used in analysis	N/A
Core Gene Group	DNA polymerase Family b exonuclease	60R	Used in analysis	N/A
Core Gene Group	DNA-dep RNA pol-II second largest subunit	62L	Used in analysis	N/A
Core Gene Group	Erv1/Alr family	88R	Used in analysis	N/A
Core Gene Group	Helicase family	21L	Excluded based on phylogeny	Yes
Core Gene Group	Hypothetical protein-Clostridium tetani	94L	Used in analysis	N/A
Core Gene Group	Immediate early protein ICP-46	91R	Used in analysis	Yes
Core Gene Group	Major capsid protein	90R	Used in analysis	N/A
Core Gene Group	Myristilated membrane protein	2L	Used in analysis	N/A
Core Gene Group	Myristilated membrane protein	53R	Used in analysis	N/A
Core Gene Group	NIF-NLI interacting factor	37R	Excluded based on phylogeny	Yes
Core Gene Group	Proliferating cell nuclear antigen	84R	Excluded based on phylogeny	N/A
Core Gene Group	Putative NTPase I	9L	Used in analysis	Yes
Core Gene Group	Putative replication factor and/or DNA binding-packing	1R	Used in analysis	N/A
Core Gene Group	Putative tyrosine kinase/lipopolysa	27R	Used in analysis	Yes

Table S2.2: Inclusion/exclusion and recombination status of genes examined in study.

	ccharide modifying enzyme			
Core Gene Group	Putative XPPG-RAD2-type nuclease	95R	Excluded based on phylogeny	N/A
Core Gene Group	Ribonuclease III	80L	Used in analysis	Yes
Core Gene Group	Ribonucleotide reductase small subunit	67L	Used in analysis	N/A
Core Gene Group	Serine-threonine protein kinase	57R	Excluded based on phylogeny	N/A
Core Gene Group	Serine-threonine protein kinase	19R	Excluded based on phylogeny	N/A
Core Gene Group	Transcription elongation factor TFIIS	81R	Used in analysis	N/A
Core Gene Group	Unknown	41R	Used in analysis	Yes
Core Gene Group	Unknown	12L	Used in analysis	N/A
RV Gene Group	ATPase-dependent protease	79R	Used in analysis	Yes
RV Gene Group	ATV p-79	89R	Used in analysis	Yes
RV Gene Group	ATV-p78	48L	Used in analysis	Yes
RV Gene Group	CARD like capsase	64R	Excluded based on phylogeny	N/A
RV Gene Group	dUTPase	63R	Excluded based on phylogeny	N/A
RV Gene Group	FV-3 orf 76R	76R	Excluded based on phylogeny	N/A
RV Gene Group	ICP-46-18	82R	Used in analysis	Yes
RV Gene Group	SGIV 0RF043R	17L	Used in analysis	N/A
RV Gene Group	SGIV-ORF009L	34R	Used in analysis	N/A
RV Gene Group	Unknown	4R	Excluded based on phylogeny	N/A
RV Gene Group	Unknown	10R	Excluded based on phylogeny	N/A
RV Gene	Unknown	11R	Excluded based	N/A

Table S2.2 (continued): Inclusion/exclusion and recombination status of genes examined in study.

Group			on phylogeny	
RV Gene Group	Unknown	14R	Excluded based on phylogeny	N/A
RV Gene Group	Unknown	23R, 24R	Excluded based on alignment	N/A
RV Gene Group	Unknown	31R	Used in analysis	N/A
RV Gene Group	Unknown	43.5L	Excluded based on alignment	N/A
RV Gene Group	Unknown	46L	Excluded based on phylogeny	Yes
RV Gene Group	Unknown	47L	Excluded based on phylogeny	Yes
RV Gene Group	Unknown	59L	Used in analysis	N/A
RV Gene Group	Unknown	69R	Excluded based on phylogeny	N/A
RV Gene Group	Unknown	74L	Used in analysis	Yes
RV Gene Group	Unknown	75L	Used in analysis	N/A
RV Gene Group	Unknown	70R	Used in analysis	N/A
F/A/R Gene Group	Hydrolase of the metallo-beta-lactamase superfamily	39R	Used in analysis	N/A
F/A/R Gene Group	Truncated eIF-2 alpha-like protein	26R	Used in analysis	N/A
F/A/R Gene Group	Unknown	40R	Used in analysis	N/A
F/A/R Gene Group	Unknown	56R	Used in analysis	N/A
F/A/R Gene Group	Unknown	71R	Used in analysis	N/A
F/A/R Gene Group	Unknown	93L	Used in analysis	N/A
Other Gene Group	Acetly-CoA Hydrolase	51R	Used in analysis	N/A
Other Gene Group	ATV p-79B	49L, 50L	Excluded based on alignment	N/A
Other Gene Group	Collagen-like	65L, 66L	Excluded based on alignment	N/A
Other Gene Group	Cytosine DNA methyltransferase (DMTase)	83R	Used in analysis	N/A

Table S2.2 (continued): Inclusion/exclusion and recombination status of genes examined in study.

Other Gene Group	Helicase B	55L	Used in analysis	N/A
Other Gene Group	NTPase-helicase like	73L	Used in analysis	Yes
Other Gene Group	Putative myeloid cell leukemia protein	97R	Used in analysis	N/A
Other Gene Group	Ribonucleotide reductase alpha subunit	38R	Used in analysis	N/A
Other Gene Group	Serine-threonine protein kinase	2.5L	Excluded based on phylogeny	Yes
Other Gene Group	Unknown	87L	Used in analysis	Yes
Other Gene Group	Unknown	77L	Used in analysis	N/A
Other Gene Group	Unknown	58.5R	Excluded based on phylogeny	N/A
Other Gene Group	Unknown	3R	Used in analysis	Yes
Other Gene Group	Unknown	25R	Excluded based on phylogeny	N/A
Other Gene Group	Unknown	28R	Excluded based on phylogeny	N/A
Other Gene Group	Unknown	33R	Excluded based on phylogeny	N/A
Other Gene Group	Unknown	45L	Used in analysis	N/A
Other Gene Group	Unknown	32R	Excluded based on phylogeny	N/A
Other Gene Group	Unknown	20R	Excluded based on phylogeny	Yes
Other Gene Group	Unknown	96R	Used in analysis	N/A

Table S2.2 (continued): Inclusion/exclusion and recombination status of genes examined in study.

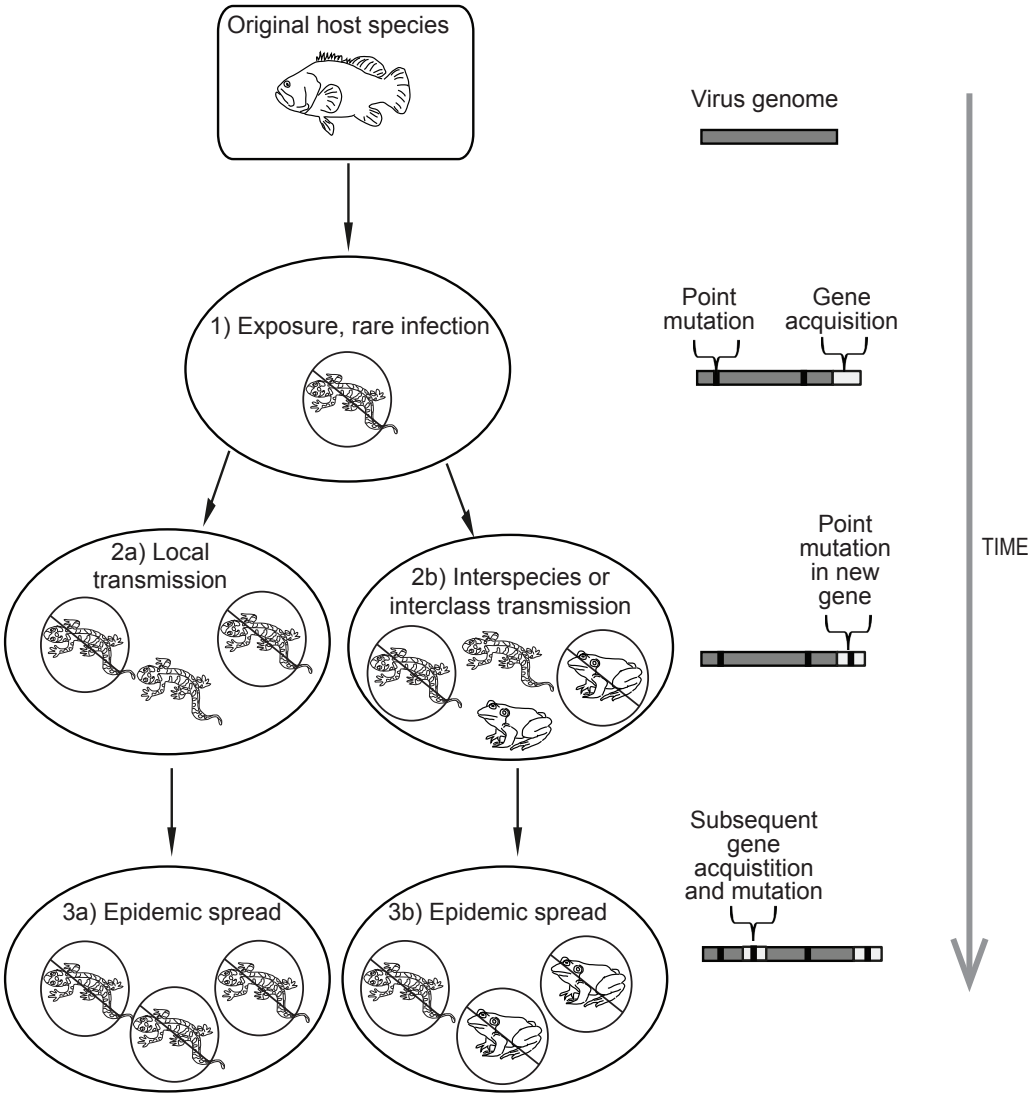

Figure 2.1: A model for Ranavirus host switching within and between vertebrate classes. During cross-species transmission, evolution of a viral genome makes the virus increasingly compatible with the genetic background of new host species, and may facilitate interclass transmission. Mutations (black bars) and the acquisition of new genes (white rectangles) may work to progressively increase the compatibility between the Ranavirus genome and new host species. Two separate scenarios that include point mutations occurring throughout the original and newly acquired regions of the genome can lead to transmission within species (2a) or between vertebrate classes (2b) over time. The acquisition of novel genes in the Ranavirus genome, and subsequent mutations within those genes, most likely plays a fundamental role in the Ranavirus host switching process.

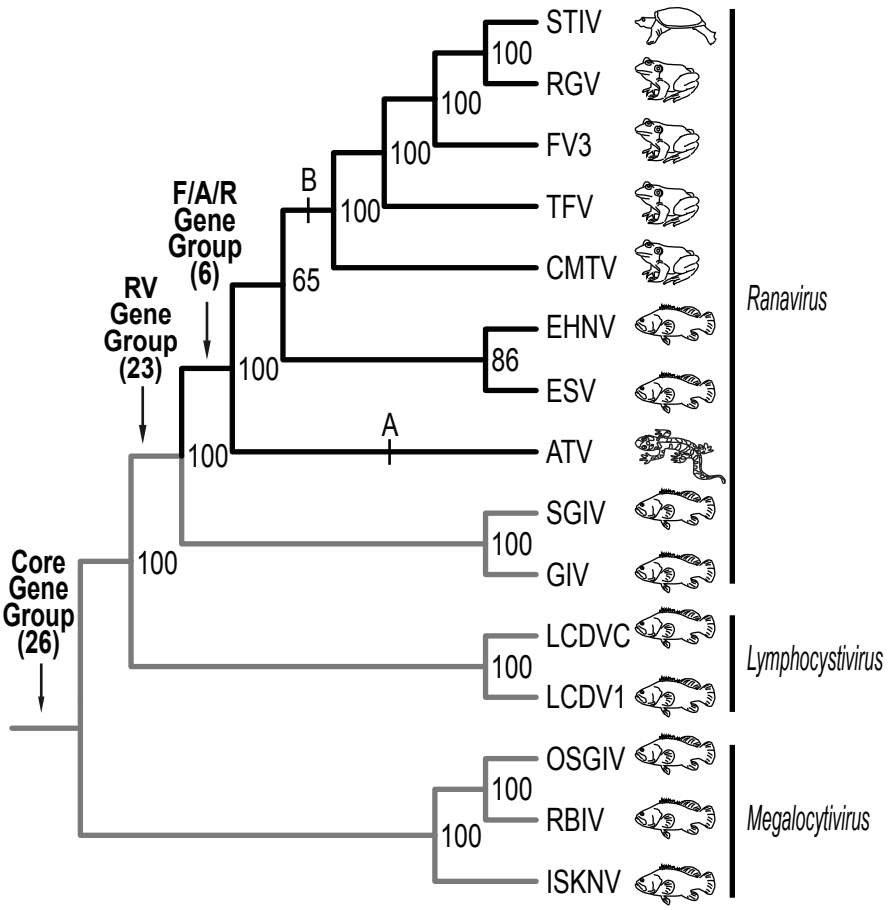

Figure 2.2: Phylogenetic relationships among vertebrate Iridoviridae genomes included in the study. A phylogenetic tree of viral species/strains isolated from fish, amphibians, and reptiles based on a maximum likelihood analysis of a concatenated data set of 26 core *Iridoviridae* genes. The Target Clade within *Ranavirus* is highlighted in black. Arrows indicate the origin of genes shared among all Iridoviridae genomes (Core gene group), among all ranavirus genomes (RV gene group), and among the Target Clade (F/A/R gene group). Tick marks indicate branches along which the initial host switch from fish to salamanders (point A) and the host switch to frogs and reptiles (point B) are thought to have occurred (Jancovich *et al*. 2010). Bootstrap values are shown. Note that each fish, amphibian, and reptile symbol denotes multiple species as indicated in Table 1.

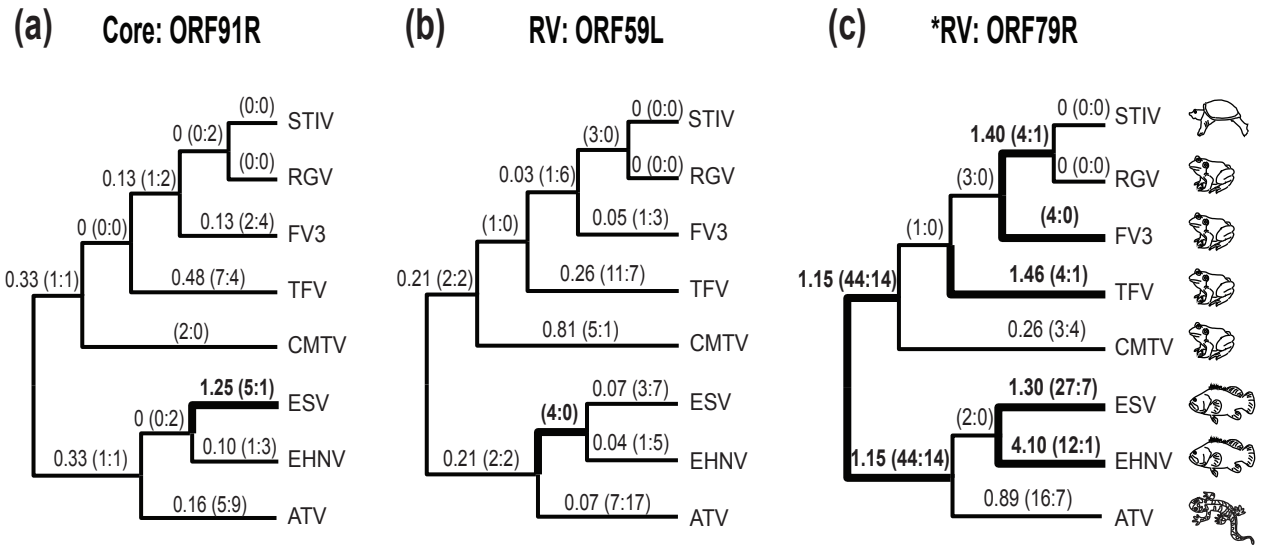

Figure 2.3: Three Core and RV genes with elevated *dN/dS* identified using branch-based tests. Phylogenies with free-ratio results for genes with branches exhibiting signatures of *dN/dS*>1 are presented. Branch values of dN/dS are shown, along with the estimated numbers of non-synonymous and synonymous (*N:S*) mutations that are predicted to have occurred along each branch (in parentheses). Branches with *dN/dS* values greater than 1 are shown in bold. Branches with incalculable dN/dS values (due to *S*=0) are highlighted in bold if the ratio of *N:S* is greater than or equal to 4:0. Genes with an asterisk were also found to exhibit signatures of positive selection under site-based tests (Table 2).

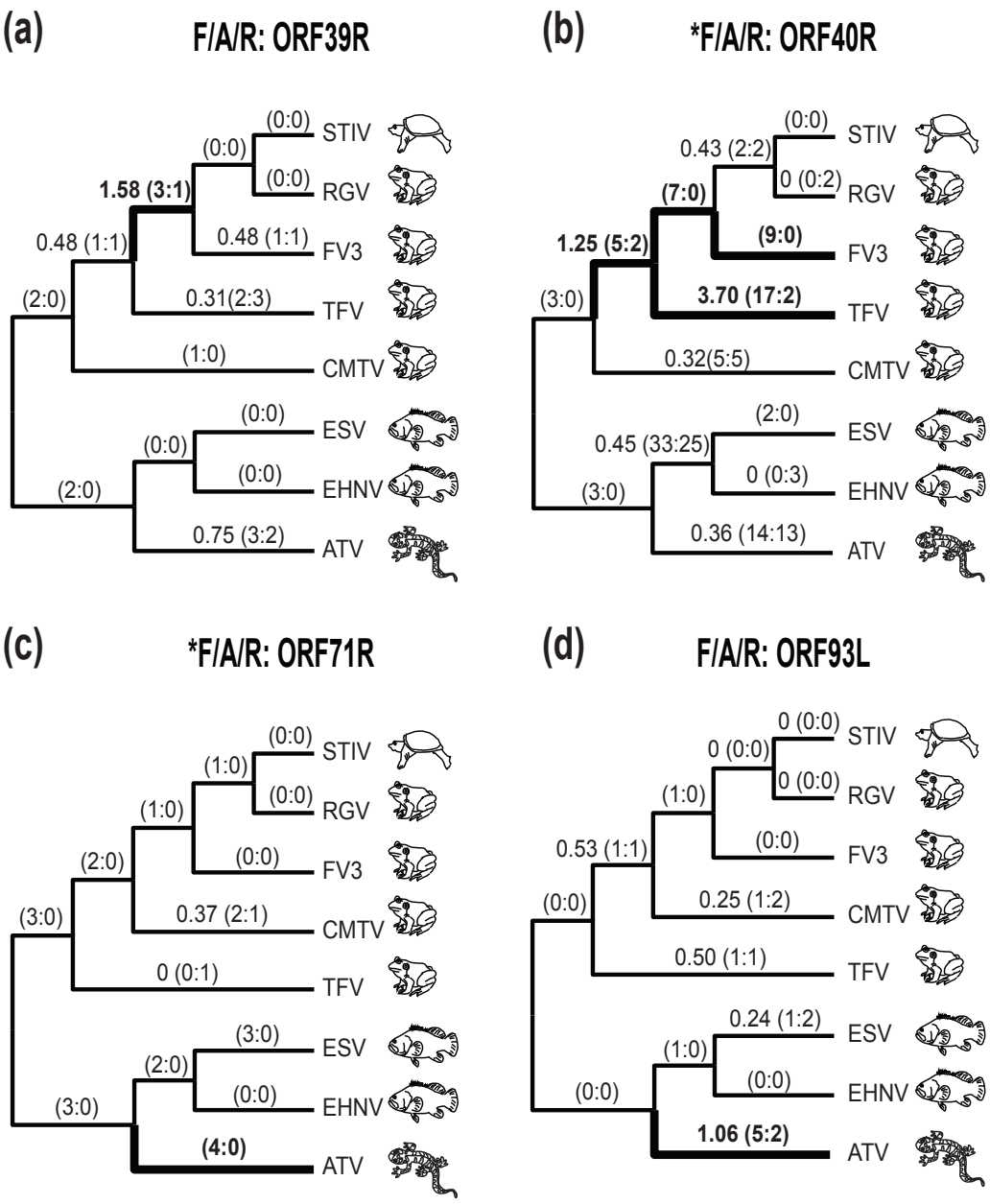

Figure 2.4: Four F/A/R genes with elevated dN/dS identified using branch-based tests. Phylogenies with free-ratio results for genes with branches exhibiting signatures of *dN/dS*>1 are presented. Branch values *of dN/dS* are shown, along with the estimated numbers of non-synonymous and synonymous (*N:S*) mutations that are predicted to have occurred along each branch (in parentheses). Branches with *dN/dS* values greater than 1 are shown in bold. Branches with incalculable *dN/dS* values (due to *S*=0) are highlighted in bold if the ratio of *N:S* is greater than or equal to 4:0. Genes with an asterisk were also found to exhibit signatures of positive selection under site-based tests (Table 2). Note variation in the topologies of ORF71R and ORF93L due to slightly different gene trees.

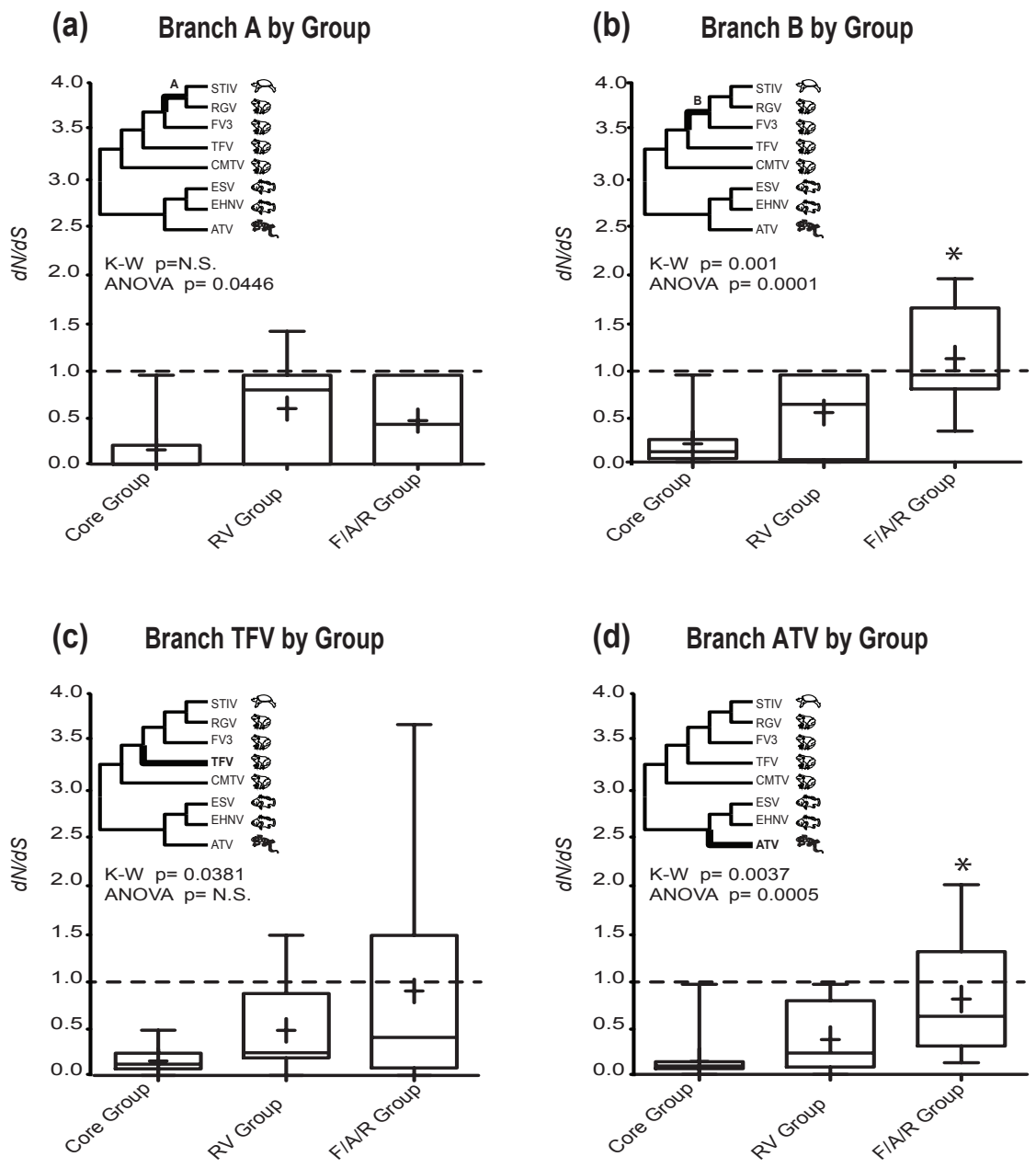

Figure 2.5: Significant differences in median and mean *dN/dS* values of Core, RV, and F/A/R gene groups detected using branch-based tests. The inset phylogenies show the significant branches in bold. Boxplots for statistically significant branches are shown (branches without significant support are not shown). Kruskal-Wallis tests (medians) were performed on the transformed data set (square root of *dN/dS*, raw data shown). One-way ANOVA tests (means) were performed on the transformed data set (square root of *dN/dS*, raw data shown). An asterisk indicates that the mean of the F/A/R gene group was statistically different from the Core gene group based on Bonferroni post-tests.

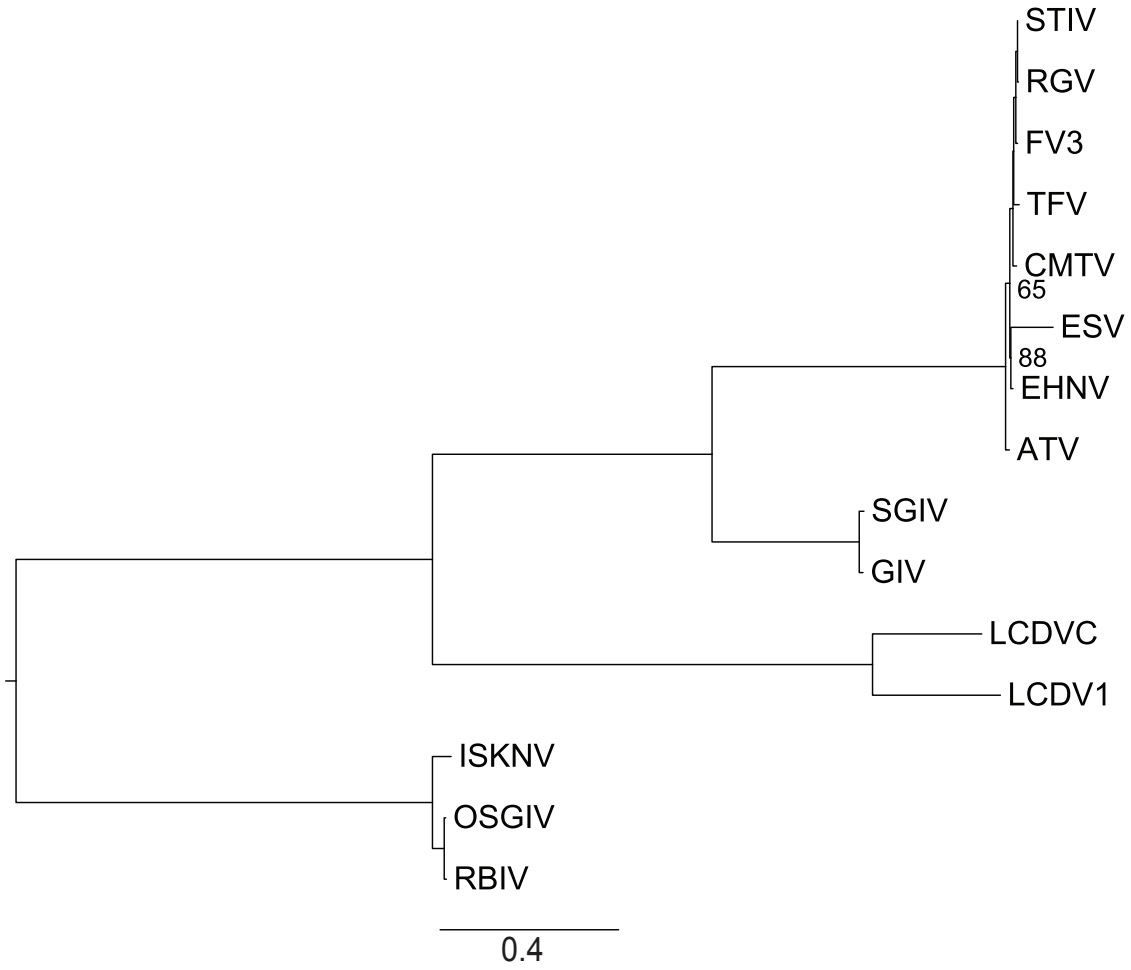

Figure S2.1: Phylogenetic relationships among vertebrate *Iridoviridae* genomes included in the study. A phylogenetic tree of viral species/strains isolated from fish, amphibians, and reptiles based on a maximum likelihood analysis of a concatenated data set of 26 core *Iridoviridae* genes. Bootstrap values less than 100 are shown.

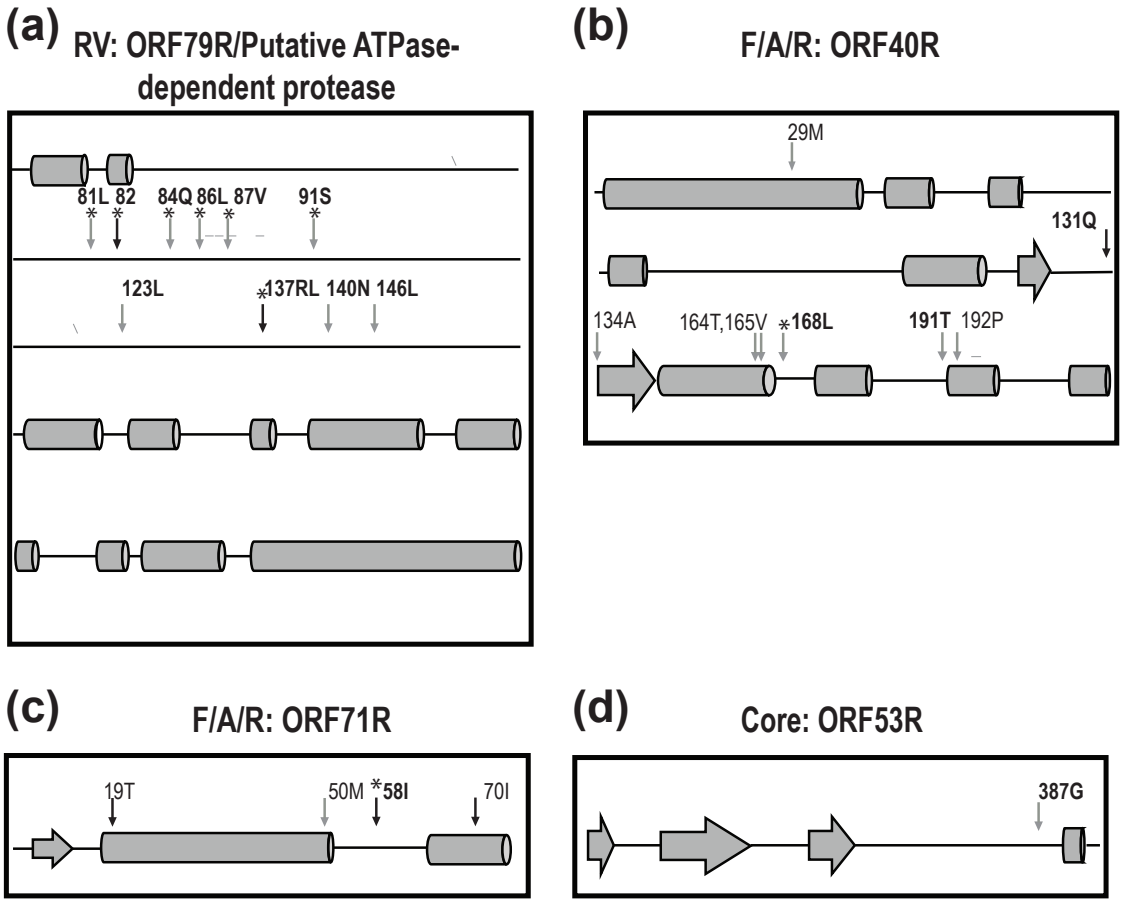

Figure S2.2: Secondary structures for proteins encoded by genes under positive selection (site-based method) show that several positively selected sites lie in unordered regions of the protein. Predicted secondary structures for ORF79R/Putative ATPase-dependent protease (a), ORF40R (b), and ORF71R (c), and ORF53R (d) are shown. α-helices (cylinders), β-strands (horizontal arrows), and unstructured regions (solid lines) are illustrated. Vertical gray arrows indicated sites with PP 0.80-0.90, and vertical black arrows indicated sites with PP> 0.90. Vertical arrows with an asterisk indicate residues in areas of high disorder confidence.

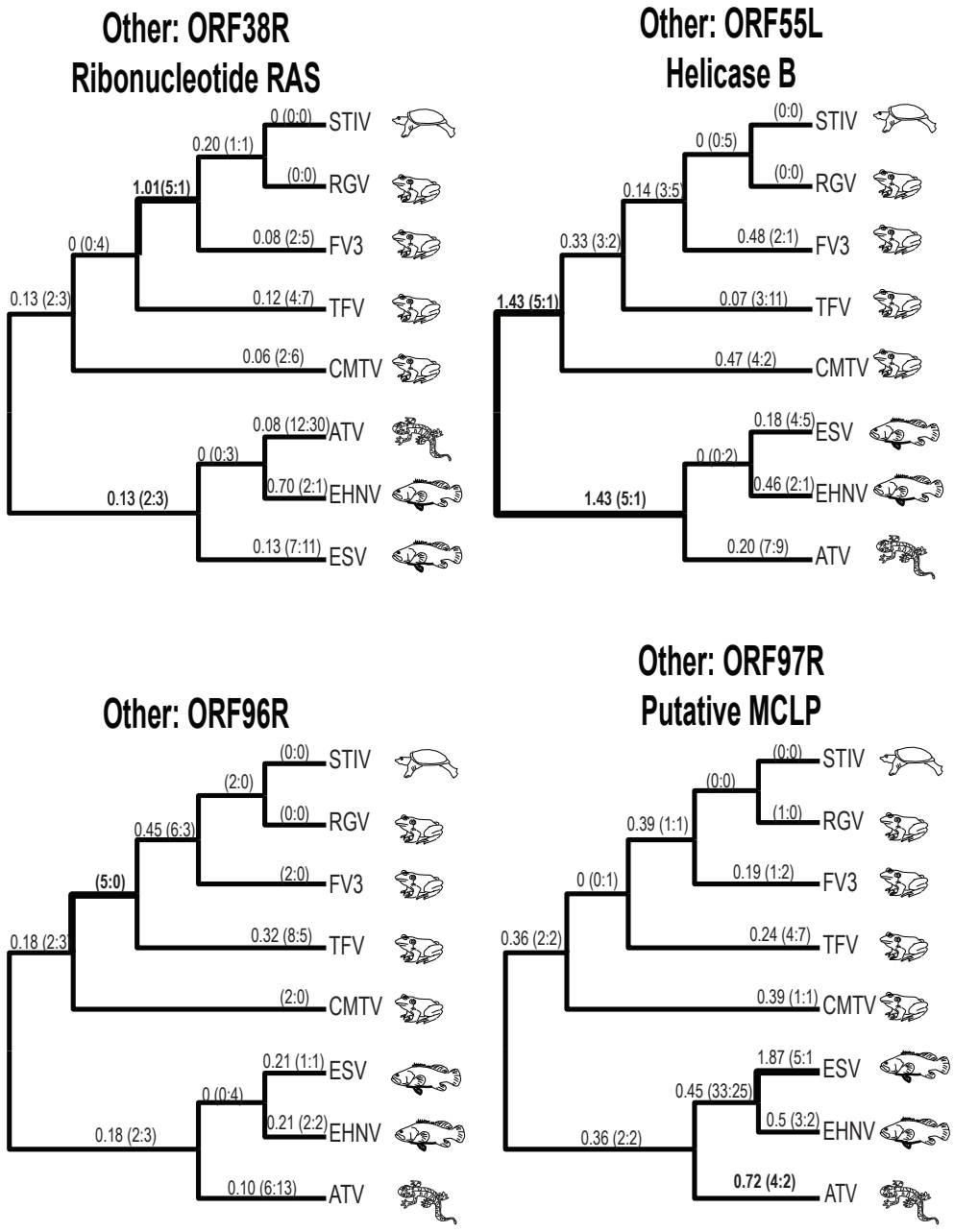

Figure S2.3: Four "Other" genes with elevated *dN*/*dS* identified using branch-based tests. Phylogenies with free-ratio results for genes with branches exhibiting signatures of *dN*/*dS*>1 are presented. Branch values of *dN*/*dS* are shown, along with the estimated numbers of non-synonymous and synonymous (*N:S*) mutations that are predicted to have occurred along each branch (in parentheses). Branches with *dN*/*dS* values greater than 1 are shown in bold. Branches with incalculable *dN*/*dS* values (due to *dS*=0) are, arbitrarily, highlighted in bold if the ratio of *N:S* is greater than or equal to 4:0. Note variation in the topology of ORF38R due to slightly different gene trees.

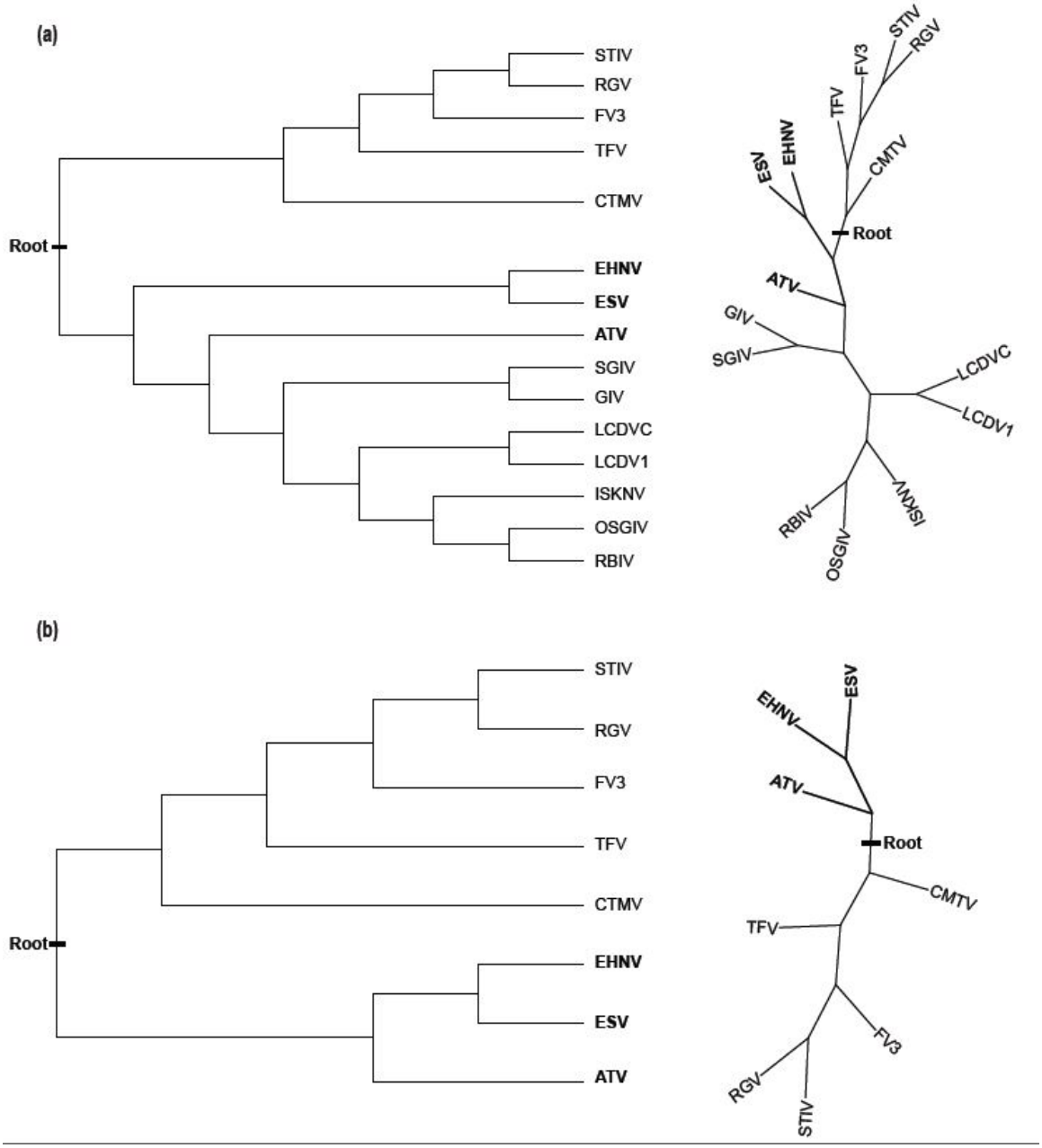

Figure S2.4: Congruent topologies between 15-taxon and 8-taxon phylogenies. a) Rooted and unrooted 15-taxon phylogenies. Branches in bold highlight the relationship between ATV, EHNV, and ESV when the root of the tree is matched to that of the 8-taxon tree. b) Rooted and unrooted 8-taxon phylogenies. Branches in bold highlight the relationship between ATV, EHNV, and ESV when the root of the tree is matched to that of the 15-taxon tree. Both unrooted and rooted trees illustrate that the (ATV, (EHNV, ESV)) relationship is consistent among the 15-taxon and 8-taxon trees.

Chapter 3: Genomic characterization of Largemouth Bass Virus (LMBV) provides further insights into *Ranavirus* taxonomic and phylogenetic relationships

ABSTRACT

Largemouth Bass Virus (LMBV) is a ranavirus that was initially isolated from *Micropterus salmoides*, and it has been identified as the etiological agent responsible for die-off events affecting both wild and cultured largemouth bass throughout the southeastern United States. Although LMBV and other closely related viruses are currently classified as ranaviruses, there has been some controversy regarding their placement within the *Ranavirus* genus. To determine the phylogenetic relationships of LMBV, the nearly complete genomic sequence of the virus was sequenced, annotated, and analyzed using standard genomic and phylogenetic techniques. The analysis of the LMBV genome suggests that it is more closely related to amphibian-like ranaviruses (ALRVs) than to grouper ranaviruses, and this is further supported by greater genomic collinearity between LMBV and ALRVs. The phylogenetic evidence supports the placement of LMBV as more closely related to grouper ranaviruses than to ALRVs. These data, in combination with other taxonomic evidence, suggest that the classification of LMBV as a ranavirus is warranted, and that LMBV and its close relatives should be considered one of three distinct *Ranavirus* subspecies.

INTRODUCTION

The *Iridoviridae* family consists of large dsDNA-containing nucleo-cytoplasmic viruses (160-200 nm) divided into five genera. These are known pathogens of a variety of invertebrate and vertebrate hosts. Three genera, *Ranavirus, Lymphocystivirus,* and

Megalocytivirus are pathogens of ectothermic vertebrates (Williams *et al*. 2005), and several of these viruses are recent emerging pathogens of commercially important cultured and wild bony fish (Whittington *et al*. 2010). Ranaviruses have received significant attention due to their role as a major factor in amphibian epizootics and population declines (Miller *et al*. 2011), but they have also had detrimental effects on fish populations. The discovery of epizootic haemotopoietic necrosis virus (EHNV) in 1985 was the first instance of an epizootic in bony fish caused by a ranavirus (Langdon *et al*. 1986). Since then, several ranavirus species and strains have been isolated from fish species worldwide (Whittington *et al*. 2010).

Largemouth Bass Virus (LMBV) is a ranavirus that was initially isolated from largemouth bass (*Micropterus salmoides*) and black crappie (*Pomoxis nigromaculatus*) in Lake Weir, Florida, and was subsequently identified as the etiological agent responsible for a largemouth bass die-off at the Santee-Cooper Reservoir in South Carolina (Plumb *et al*. 1996; Grizzle *et al*. 2002). Since the first reported cases, systematic surveys of wild and cultured populations of largemouth bass have identified the pathogen across the southeastern United States (Plumb *et al*. 1999; Hanson *et al*. 2001; Maceina and Grizzle 2006; Neal *et al*. 2009; Southard *et al*. 2009; Blazer *et al*. 2010; National Wildlife Health Survey 2013). Based on molecular characterizations, LMBV is currently grouped with two other viruses, Doctor Fish Virus (DFV) and Guppy Virus 6 (GV6), in a group collectively referred to as the Santee-Cooper ranaviruses. Although DFV and GV6 are identical within the major capsid protein (MCP), the MCP sequence for LMBV differs slightly (Ohlemeyer *et al*. 2011). Additionally, DFV infection in largemouth bass results in ulceration and hemorrhaging of the body surface (Deng *et al*. 2011), while LMBV infection of largemouth bass leads to over-inflation of the swim bladder but no external lesions (Plumb *et al*. 1999). Both DFV and GV6 were isolated from the ornamental fish

Labriodes dimidatus and *Poecilia reticulata,* respectively, which were imported from Southeast Asia (Mao *et al*. 1999, Ohlemeyer *et al*. 2011), which suggests that geographically restricted LMBV may have evolved from a virus that was introduced to the United States through the pet trade. This hypothesis is further supported by the isolation and characterization of a virus collected from cultured largemouth bass in China that was identical to DFV (based on the MCP sequence), which illustrates that largemouth bass are susceptible to DFV-like viruses (Deng *et al*. 2011).

The International Committee on the Taxonomy of Viruses (ICTV) currently lists six *Ranavirus* species: Frog Virus 3 (FV3), *Ambystoma tigrinum* virus (ATV), Bohle Iridovirus (BIV), EHNV, European Catfish Virus (ECV), and Santee-Cooper Ranavirus. Ranavirus species are distinguished by unique restriction fragment length polymorphism profiles, distinct viral protein profiles, <95% sequence identity/similarity between select viral genes, and a distinctive range of susceptible host species (King *et al*. 2012). Early examinations of molecular characteristics resulted in the classification of LMBV as a ranavirus (Mao *et al*. 1999). However, a subsequent study of the Santee-Cooper ranaviruses DFV and GV6 led Hyatt *et al*. (2000) to conclude that these viruses were not members of *Ranavirus* based on low hybridization with FV3, but that they represented a separate but related group of viruses. Subsequent genomic analyses suggested the formation of two *Ranavirus* subspecies, the "grouper" ranaviruses (which includes Singapore grouper iridovirus [SGIV] and Grouper iridovirus [GIV]), and "non-grouper ranaviruses," also referred to as amphibian-like ranaviruses (ALRVs) (Eaton *et al*. 2007; Jancovich *et al*. 2010). However, the genomic study did not include Santee-Cooper Ranavirus genomic data. More recently, the complete MCP sequences of the three Santee-Cooper ranaviruses were obtained, and analyses showed a 70% identity to the MCP sequences of grouper ranaviruses and 78% identity to that of ALRVs (Ohlemeyer *et*

al. 2011). Using this information, Ohlemeyer *et al*. (2011) suggested the breakdown of *Ranavirus* into three distinct subspecies: grouper ranaviruses, Santee-Cooper ranaviruses, and ALRVs. However, these suggestions have not been addressed by the ICTV and the taxonomic position of Santee-Cooper Ranavirus is still in question.

In addition to the taxonomy of ranaviruses, the occurrence of host switching within the genus is also of great interest. *Ranavirus* is unique among the *Iridoviridae* viruses isolated from vertebrate hosts in that both *Lymphocystivirus* and *Megalocytivirus* solely infect fish species, while ranaviruses are known pathogens of fish, amphibians, and reptiles. The extensive host range of ranaviruses is presumably due to recent host switching events. Jancovich *et al*. (2010) used available genomic data to hypothesize two potential ranavirus host-switching patterns. The first suggests two separate jumps from 1) fish to salamanders and 2) fish to frogs, and a subsequent jump from frogs to turtles. The second scenario includes a single jump from fish to salamanders/frogs with a subsequent jump from frogs to turtles, and a secondary switch from salamanders back to fish. Abrams *et al*. *(*2013), using additional genomic data, supported a secondary jump from salamanders back to fish. However, the genomic data did not include sequences from Santee-Cooper ranaviruses, which could aid the resolution of the moderately supported nodes.

Here, we present the nearly complete *de novo* genomic sequence of the Santee-Cooper Ranavirus, LMBV, which was attained via next-generation sequencing. The goals of this study were to obtain the gene complement of LMBV to determine its position in the *Iridoviridae* phylogeny, to provide insight into the contested phylogenetic position of LMBV within *Ranavirus*, and to clarify the sequence of ranavirus host switching events. The available genomic content suggests that LMBV is more closely related to ALRVs than grouper ranaviruses, and this is further supported by greater genomic collinearity

between LMBV and ALRVs. However, the limited genomic collinearity identified between LMBV and all other ranaviruses supports the intermediate position of LMBV within the genus. The phylogenetic evidence also supports the placement of LMBV as more closely related to ALRVs than grouper ranaviruses. The addition of LMBV sequence data provides strong support for two major host switches from fish to salamanders/frogs and frogs to turtles; thus eliminating the hypothesis of a secondary switch from salamanders back into fish. These data, in combination with other taxonomic evidence, suggest that the classification of LMBV as a ranavirus is warranted, and that Santee-Cooper ranaviruses should be considered one of three distinct *Ranavirus* species.

MATERIALS AND METHODS

Sequencing and Assembly

LMBV DNA was isolated from viral stocks that were originally collected from diseased *M. salmoides* from the Santee Cooper Reservoir in South Carolina, which is the site of one of the first LMBV-related die-off events (Grizzle *et al*. 2002). The purified DNA sample was provided by Dr. James Jancovich of California State University at San Marcos. Purified viral DNA (100 ng/μl) was then sheared to create a long-insert paired-end library that was subsequently subjected to high-throughput sequencing on an Illumina HiSeq2000 instrument housed at the Genomic Sequencing and Analysis Facility at the University of Texas at Austin. The paired-end output included approximately 12 million sequences.

The *de novo* assembly of the sequence data was generated using Velvet version 1.2.03 (Zerbino and Birney 2008) with a hash length (k-mer value) of 73, a coverage

cutoff of 10, and an insert length of 700; the analysis yielded 171 contigs. The resulting assembly was then used as the input for Glimmer: Gene Locator and Interpolated Markov ModelER version 3.02 (Delcher *et al*. 1999), which was used to identify 183 predicted open reading frames (ORFs). Several other *de novo* assembly parameters were tested; however, the results from the parameters reported here were chosen because they provided the fewest and longest predicted ORFs.

Genome Annotation

The predicted ORFs were then manually compared to the protein sequences stored in the National Center for Biotechnology Information database using the BLASTX similarity search algorithm (http://blast.ncbi.nlm.nih.gov). All predicted ORFs were analyzed regardless of length. The predicted function, best BLASTX hit with corresponding accession number, percent identification, expect value, percent positive, and percent gap values were recorded. When applicable, each ORF was assigned to the "Core" gene group, the "RV" (ranavirus*)* gene group, the "F/A/R" (Fish/Amphibian/Reptile) gene group, or the "Other" gene group. The "Core" gene group includes 26 genes found in all completely sequenced *Iridoviridae* genomes. The "RV" gene group comprises 23 genes specific to all completely sequenced ranavirus genomes. The F/A/R gene group includes 6 genes found exclusively in the completely sequenced ALRVs, which is the monophyletic group of ranaviruses, excluding Singapore Grouper Iridovirus (SGIV) and Grouper Iridovirus (GIV). Lastly, the "Other" gene group is comprised of an additional 20 genes, which includes genes found in the completely sequenced genomes of ranavirus and at least one other *Iridoviridae* genus (i.e., *Ranavirus* and *Chloriridovirus* but no other genera). The gene groups are based on the genomic analysis of Eaton *et al*. (2007), which identified orthologous genes throughout the

Iridoviridae family using a combination of BLAST searches and Viral Orthologous Clusters software.

In several cases, multiple predicted ORFs matched the same previously identified *Iridoviridae* gene or conserved domain. Each predicted ORF sequence was aligned to the target sequence from the *Iridoviridae* genome with the best BLASTX expect value to determine the quality of the alignment and the placement along the target sequence. In cases where multiple LMBV ORFs aligned to consecutive regions of the target sequence, the sequences were concatenated and reanalyzed as a single predicted ORF, and the BLASTX results were reported. Predicted ORFs that were poorly aligned to the target sequence were excluded from further analyses. Twenty predicted ORFs did not match any known gene or conserved domain based on the BLASTX results, so the sequences were reanalyzed using BLASTN and TBLASTX similarity algorithms, yet no matches were identified. Five predicted ORFs, exceeding a length of 200 bp, that were not matched to any known gene or conserved domain were considered putative unique LMBV ORFs; all sequences shorter than 200 bp were excluded. Moreover, 44 predicted ORFs that only matched cellular sequences or conserved domains were considered cellular contamination and were also excluded. Based on the known genome sizes of other ranaviruses, it was apparent that the 87 predicted LMBV ORFs that were identified did not represent the complete LMBV genome. Therefore, mapping of the predicted ORFs to determine the gene overlap and transcriptional orientation was not conducted.

Phylogenetic and Dot Plot Analyses

To determine the phylogenetic placement of LMBV within the vertebrate *Iridoviridae* viruses, a phylogenetic tree was reconstructed via maximum likelihood using a concatenated dataset of the 26 Core group genes (Table 3.1). The complete genomes

used in this analysis are listed in Table 3.2. The sequences for each gene were aligned using Clustal W (Larkin *et al*. 2007), and the 26 alignments were subsequently concatenated. The phylogeny was reconstructed using RAxML (Stamatakis 2006) under the GTR + Γ model with the data partitioned by gene, and 500 bootstrap replicates were used to assess node support (Figure 3.3).

Dot plot analyses were conducted using the JDotter program (Brodie *et al*. 2004) to identify regions of similarity between the LMBV genomic sequence and the complete genomic sequences of SGIV (isolated from a marine fish host), EHNV (isolated from a freshwater fish host), ATV (isolated from a salamander host), and FV3 (isolated from a frog host) ranaviruses. Regions of similarity between the genomes are indicated by the presence of a straight line (Figure 3.2).

RESULTS

87 Predicted LMBV ORFs Identified Via *De Novo* Sequencing Analysis

The *de novo* genome assembly using the Velvet and Glimmer programs identified a total of 183 predicted open reading frames (ORFs). Of the 183 predicted ORFs, several represented sections of the same gene, 44 ORFs were classified as cellular contamination, and an additional 20 ORFs were excluded based on the lack of matching homologues and short sequence length (<200 bp). The BLASTX analyses of the predicted LMBV ORFs identified 26 of the 26 genes (100%) from the Core gene group, 19 of the 23 genes (83%) from the RV gene group, 2 of the 6 genes (33%) from the F/A/R gene group, and 19 of the 20 genes (95%) from the Other gene group (Table 3.3). In addition, 15 predicted LMBV ORFs were not assigned to either of the above gene groups, but the ORFs did match previously identified ranavirus genes. The predicted LMBV ORF 126 was similar

to the K2R protein from poxviruses, which contain a tumor necrosis factor receptor (TNFR) domain. The similarity to poxvirus genes is not unexpected since many of the known *Iridoviridae* homologs were first identified and characterized in poxviruses, which are also large dsDNA-containing nucleo-cytoplasmic viruses. Lastly, an additional 5 predicted ORFs with lengths exceeding 200 bp were tentatively labeled as unique LMBV ORFs. In total, 87 predicted LMBV ORFs ranging in size from 35.1 kDa to 1,215.5 kDa were annotated.

The best matches for the BLASTX analyses returned hits for each of the completely sequenced ranavirus genomes (Figure 3.1A). Moreover, a closer look at the LMBV ORFs that best-matched previously identified ranavirus genes indicted that the majority (80.25%) of the predicted ORFs more closely matched genes from ALRVs than those from grouper ranaviruses (Figures 3.1B and C). This suggests a higher degree of similarity between LMBV and the ALRVs. The number of genes found in completely sequenced ranavirus genomes averages around 100 genes for ALRVs and approximately 140 genes for grouper ranaviruses. Although the number of predicted ORFs identified here is lower than these expected values, the higher degree of similarity with ALRVs suggests that the size of the complete LMBV genome may exhibit an intermediate size. Moreover, since the LMBV genomic sequence is incomplete, it is possible that the predicted ORFs presented only represent a specific region of the genome. To determine if the predicted LMBV ORFs were localized or evenly distributed across the genome, the genomic map of EHNV (Jancovich *et al*. 2010) was used as a template to map the location of the predicted LMBV ORFs. The results of this mapping indicated that the predicted LMBV ORFs are evenly distributed across the EHNV genomic map (data not shown).

Dot plot analyses show limited regions of collinearity when compared to SGIV, EHNV, ATV, and FV3 genomes

To further determine the relationship between LMBV and other ranaviruses, dot plot analyses were used to assess the levels of collinearity between the LMBV sequence and the complete genomic sequences of SGIV (isolated from a marine fish host), EHNV (isolated from a freshwater fish host), ATV (isolated from a salamander host), and FV3 (isolated from a frog host). A -45° line is expected when two genomic sequences exhibit complete collinearity, and breaks in this line are representative of sequences found in one genome but not the other. Previous dot plot analyses indicated high levels of collinearity between ALRV genomes; however limited areas of collinearity were detected when grouper ranavirus genomes were compared to those of ALRVs (Jancovich *et al*. 2010; Lei *et al*. 2012; Mavian *et al*. 2012a). The dot plots comparing the LMBV sequence to the selected ranavirus genomes yielded limited regions of collinearity for all analyzed pairs (Figure 3.2). The regions of collinearity are more prevalent in comparisons of the LMBV sequence to ALRVs (29-30 identified regions) as compared to the grouper ranavirus, SGIV (19 regions of collinearity). This data supports the intermediate position of LMBV between the grouper ranaviruses and ALRVs, and suggests a closer relationship between LMBV and ALRVs. However, the complete LMBV genomic sequence is needed to fully resolve the genomic similarity.

Phylogenetic data supports a closer LMBV-ALRV relationship and identifies two major host-switches in *Ranavirus*

To determine the phylogenetic position of LMBV within the *Iridoviridae,* the phylogeny was reconstructed using a concatenated sequence of the 26 Core genes from the available *Iridoviridae* genomes (Table 3.2) and the corresponding predicted LMBV

ORFs (Table 3.1). The phylogenetic tree shows that LMBV holds an intermediate position between the grouper ranaviruses and ALRVs (Figure 3.3). Moreover, the branch leading from grouper ranaviruses to LMBV is slightly longer than the branch leading from LMBV to ALRVs; thus suggesting that LMBV is more closely related to ALRVs. The phylogeny also illustrates the relatively rapid evolution within the ALRV clade, where the most recent and extensive host switching has taken place.

The addition of the LMBV sequence broke the long branch separating the grouper and ALRV clades, and added greater resolution to the ranavirus phylogeny. As a result, the region of the phylogeny that previously supported a secondary switch from salamanders back to fish (Abrams *et al*. 2013) is now resolved to show a single, clear transition from fish to salamanders with strong bootstrap support (Figure 3.4). Therefore, the phylogeny implies two major host switches: an initial jump from fish to salamanders/frogs and a more recent jump from frogs to turtles.

DISCUSSION

The identification and analysis of the 87 putative LMBV ORFs not only advance our understanding of *Ranavirus* taxonomy, but also refine our knowledge of host switching within the genus. Initial studies of individual LMBV genes and other molecular characteristics highlighted the similarity of LMBV to other ranaviruses (Mao *et al*. 1999). However, documentation of differences led to questioning of the taxonomic placement of LMBV and other Santee-Cooper ranaviruses within the genus (Hyatt *et al*. 2000). The identification of homologs of ranavirus-specific genes (83% of RV genes and 33% of F/A/R genes) suggests that LMBV should be classified as a ranavirus. The intermediate position of LMBV between the grouper ranaviruses and ALRVs is

supported by the limited regions of collinearity observed in the dot plots (Figure 3.2) and the position of LMBV in the *Iridoviridae* phylogeny (Figure 3.3). These data, in conjunction with other molecular data (Mao *et al*. 1999; Hyatt *et al*. 2000; Ohlemeyer *et al*. 2011), support the recommendation of Ohlemyer *et al*. (2011) that *Ranavirus* species should be classified as one of three distinct subspecies: grouper ranaviruses, Santee-Cooper ranaviruses, and ALRVs.

The taxonomic classification of ranaviruses was originally based on biological properties of the viruses, but the more recent inclusion of sequence data has enhanced the ability to classify viral species and strains. However, the high degree of genetic similarity in ALRVs combined with the haphazard naming of identical and nearly identical viruses based on geographical or host species information has hampered the process by increasing the number of unclassified isolates (Chinchar *et al*. 2009). The data presented here clearly show how the addition of genomic-scale data can resolve ranavirus phylogenetic relationships, and this highlights the need for further genomic sequencing. In addition to adopting the three distinct *Ranavirus* subspecies, the ICTV should also consider implementing taxonomic rules based on phylogenetic data as outlined in the PhyloCode (Cantino and de Queiroz 2010). The combination of phylogeny-based taxonomy and traditional taxonomic data may act to simplify and expedite the viral classification process.

The inclusion of LMBV sequence data provided a clear depiction of the major ranavirus host-switching events. The phylogeny shows two major transitions: fish to salamanders/frogs and frogs to turtles (Figure 3.4). Abrams et al. (2013) identified several ranavirus-specific genes that appear to have undergone adaptive evolution, which may have facilitated interspecies and interclass host switching, and two of these genes were recovered in the LMBV genomic sequence (Table 3.3). The first gene, a putative

ATPase-dependent protease (FV3 ORF 79R) from the RV gene group, exhibited evidence of positive selection acting on several sites in an unstructured region (neither α-helix nor β-strand) of the protein. The alignment of the predicted homolog LMBV ORF 85 to the sequences from all other ranaviruses indicated that the region where positive selection is acting (approximately 100 amino acids) is absent in LMBV and grouper ranaviruses (Figure S3.1). Considering that viral proteases are thought to be essential to viral replication and infectivity (Babe & Craik 1997), the acquisition of this gene region by ALRVs may have contributed to viral ability to infect and replicate in naïve host species. The second gene, FV3 ORF 71R (F/A/R gene group), was previously thought to only occur in ALRV genomes; therefore, the occurrence of this gene in LMBV is surprising. The previous analysis of FV3 ORF 71R identified a signature of positive selection (an elevated *dN*/*dS* value) along the branch leading to the salamander virus ATV. A reanalysis of the data following Abrams *et al*. (2013) and the addition of the predicted homolog LMBV ORF 147 recovered a consistent pattern (Supplemental Figure S3.2). This suggests that even though the gene was present in ranaviruses prior to the evolution of ALRVs, positive selection may have acted on the gene after the initial jump from fish to salamanders.

The identification of LMBV homologs of the immune evasion genes eif-2α homolog (LMBV ORF 82) and the K2R protein containing a TNFR domain from poxviruses (LMBV ORF 126) is particularly interesting considering the hypothesis that LMBV evolved from a virus that was possibly introduced to the United States via the pet trade. A gene knockout study of eif-2α homolog showed that the impairment of the gene resulted in decreased viral growth and virulence (Chen *et al*. 2011); therefore, the presence of the gene significantly enhances viral fitness. Moreover, the poxvirus K2R protein is thought to function as a mechanism for the evasion of the host immune

response (Shchelkunov 2003). Considering that eif-2α homolog first appears in the LMBV genome (not present in grouper ranaviruses), and that a homolog for the K2R protein is absent in all other sequenced *Iridoviridae* genomes, it is reasonable to suggest that the genes may have played a role in the either the initial transition from marine to freshwater fish, or in the spread of Santee-Cooper-like viruses to naïve fish hosts. Goldberg *et al*. (2003) identified intraspecific strain variation in LMBV isolates that was associated with differences in virulence. Further studies of LMBV should examine population-level variation in the eif2α homolog and K2R genes to further understand the roles that the genes play in LMBV evolution and spread.

Although the complete LMBV genome sequence has not been obtained, the data presented here are a step forward in the resolution of ranavirus taxonomy and evolution. LMBV is a major pathogen of largemouth bass, which is considered an alien species in nearly 80 countries and has been named one of the world's 100 worst invaders (Global Invasive Species Database 2013). In addition to the negative impact on the ecology of the environments to which it is introduced, it is possible that largemouth bass are also vectors of LMBV or other ranaviruses (Picco *et al*. 2010). The *Iridoviridae* phylogeny indicates relatively rapid evolution in the ALRV clade where extensive interspecies and interclass host switching has taken place. LMBV represents a transitional point in the evolution of ranaviruses, and ecological and evolutionary studies of the virus may provide a unique opportunity to understand the mechanisms that contributed to the initial host switch from fish to amphibians. Therefore, it is imperative that we learn as much as possible about the evolution of these viruses and the genes that impact their fitness.

Core Gene Name	Predicted LMBV ORF	FV3 Homolog
ATPase-like protein	ORF 74	ORF 15R
Deoxynucleoside kinase	ORF 40	ORF 85R
DNA polymerase	ORF 131	ORF 60R
DNA-dependent RNA polymerase II largest subunit	ORF 113	ORF 8R
DNA-dependent RNA polymerase II second largest subunit	ORF 114	ORF 62L
ErvlAlr/ thiol oxidoreductase	ORF 35	ORF 88R
Helicase family	ORF 51	ORF 21L
Hypothetical protein-Clostridium tetani	ORF 95	ORF 94L
Immediate early protein ICP-46	ORF 136	ORF 91R
Major capsid protein	ORF 138	ORF 90R
Myristylated membrane protein	ORF 12	ORF 53R
NIF/NLI interacting factor	ORF 107	ORF 37R
Proliferating cell nuclear antigen	ORF 38	ORF 84R
Putative D5 family NTPase/ATPase	ORF 48	ORF 22R
Putative myristylated membrane protein	ORF 142	ORF 2L
Putative NTPase	ORF 88	ORF 9L
Putative replication factor and/or DNA binding-packing	ORF 163	ORF 1R
Putative tyrosine kinase	ORF 86	ORF 27R
Putative XPPG-RAD2-type nuclease	ORF 94	ORF 95R
Ribonuclease III	ORF 117	ORF 80L
Ribonucleotide reductase small subunit	ORF 159	ORF 67L
Serine-threonine protein kinase	ORF 6	ORF 19R
Serine-threonine protein kinase	ORF 103	ORF 57R
Transcription elongation factor TFIIS	ORF 121	ORF 81R
Unknown protein	ORF 44	ORF 12L
Unknown protein	ORF 72	ORF 41R

Table 3.1: Twenty-six core Iridoviridae genes and the corresponding predicted LMBV ORFs with FV3 homologs as a reference.

Viral Species/Strain	Genus	Host Species	Genbank Accession No.
Frog virus 3 (FV3)	*Ranavirus*	*Rana pipiens*	AY548484 (Tan *et al*. 2004)
Soft-shelled turtle iridovirus (STIV)	*Ranavirus*	*Trionyx sinensis*	EU627010 (Huang *et al*. 2009)
Rana grylio virus (RGV)	*Ranavirus*	*Rana grylio*	JQ654586 (Lei *et al*. 2012)
Tiger frog virus (TFV)	*Ranavirus*	*Rana tigrina*	AF389451 (He *et al*. 2001)
Common midwife toad virus (CMTV)	*Ranavirus*	*Alytes obstetricans*	JQ231222 (Mavian *et al*. 2012a)
Ambystoma tigrinum virus (ATV)	*Ranavirus*	*Ambystoma tigrinum*	AY150217 (Jancovich *et al*. 2003)
Epizootic haematopoietic necrosis virus (EHNV)	*Ranavirus*	*Oncorhynchus mykiss*	FJ433873 (Jancovich *et al*. 2010)
European sheatfish virus (ESV)	*Ranavirus*	*Silurus glanis*	JQ724856 (Mavian *et al*. 2012b)
Singapore grouper iridovirus (SGIV)	*Ranavirus*	*Epinephelus tauvina*	AY521625 (Song *et al*. 2004)
Grouper iridovirus (GIV)	*Ranavirus*	*Epinephelus awoara*	AY666015 (Tsai *et al*. 2005)
Lymphocystis disease virus 1 (LCDV1)	*Lymphocystivirus*	*Platichthys flesus*	L63545 (Tidona and Darai 1997)
Lymphocystis disease virus China (LCDVC)	*Lymphocystivirus*	*Paralichthys olivaceus*	AY380826 (Zhang *et al*. 2004)
Infectious spleen and kidney necrosis virus (ISKNV)	*Megalocytivirus*	*Siniperca chuatsi*	AF371960 (He *et al*. 2002)
Orange-spotted grouper virus (OSGIV)	*Megalocytivirus*	*Epinephelus coioides*	AY894343 (Lu *et al*. 2005)
Rock bream iridovirus (RBIV)	*Megalocytivirus*	*Oplegnathus fasciatus*	AY532606 (Do *et al*. 2004)

Table 3.2: Fifteen genomes utilized in the present study and their host species.

ORF	Size (no. of amino acids)	Mol. Weight (kDa)	Predicted Function	Gene Group	Best-matching iridovirus ORF	Expect Score	% Identity	% Positive	Gap %	GeneBank access. no.
2	180	166,866	Unknown protein	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4	566	524,163	Putative acetyl-CoA hydrolase	Other	ESV 109R	2e-174	46	66	1	AFJ52394.1
5	144	133,233	Unknown protein (membrane)	Other	ESV 116L	2e-53	62	74	4	AFJ52401.1
6[i]	898	834,259	Serine-threonine protein kinase	Core	TFV 19R	0.0	42	54	12	ABB92284.1
8	103	95,545	Putative acetyacetyl-CoA reductase	N/A	ESV 119R	2e-16	43	58	10	AFJ52405.1
12	307	284,768	Myristylated membrane protein	Core	CMTV 56L	1e-113	69	82	0	AFA44962.1
16[ii]	336	312,061	3-beta hydroxyl-steroid dehydro-genase/ isomerase	Other	CMTV 57R	2e-153	63	77	1	AFA44963.1
18[iii]	364	337,597	Unknown protein	RV	SGIV 92L	8e-08	21	44	7	AAS18107.1
19	262	243,396	p31K protein	Other	CMTV 82L	1e-132	77	89	0	AFA44988.1
29	178	165,470	Unknown protein	N/A	CMTV 45L	8e-27	58	76	0	AFA44950.1
34	260	242,859	Unknown protein	RV	CMTV 17L	8e-21	39	50	16	AFA44921.1
35	128	118,734	ErvlAlr/ thiol oxido-reductase	Core	TFV 94R	6e-60	66	82	0	ABB92342.1
36	594	550,510	Unknown protein	Other	FV3 87L	2e-83	35	52	5	AAT09747.1
38	254	235,669	Proliferating cell nuclear antigen	Core	TFV 90R	2e-91	60	74	1	ABB92338.1
40[iv]	206	191,650	Deoxy-nucleoside kinase	Core	GIV 40L	1e-33	65	74	1	AAV91061.1
41	136	126,284	Unknown protein	RV	FV3 47L	2e-21	38	52	4	AAT09706.1
43	189	175,938	Unknown protein	Other	FV3 49L	3e-52	64	83	0	AAT09708.1
44[v]	228	211,573	Unknown protein	Core	ATV 87R	4e-56	69	84	0	AAP33268.1
45	70	65,109	Unknown protein	RV	FV3 11R	3e-13	54	75	0	AAT09670.1
48[vi]	954	886,483	Putative D5 family NTPase/ ATPase	Core	FV3 22R	0.0	75	85	1	AAT09681.1

Table 3.3: Predicted LMBV open reading frames (ORFs).

50	539	499,901	Unknown protein	N/A	N/A	N/A	N/A	N/A	N/A	N/A
51	208	193,524	Helicase family	Core	STIV 24L	1e-90	72	80	3	ACF42243.1
53[vii]	416	386,057	Ribo-nucleoside-diphosphate reductase large subunit	Other	CMTV 69L	0.0	77	85	2	AFA44975.1
55	154	143,030	Unknown protein	N/A	SGIV 126R	1e-31	44	56	3	AAS18141.1
56	371	344,049	Unknown protein	RV	ATV 45L	1e-58	32	54	5	AAP33224.1
57	208	192,753	Unknown protein	N/A	CMTV 29R	2e-45	40	55	1	AFA44933.1
58	73	67,611	Unknown protein	RV	ESV 40L	2e-26	62	74	0	AFJ52323.1
59	85	78,432	Putative LITAF/PIG7 possible membrane associated motif in LPS-induced tumor necrosis factor alpha factor	RV	FV3 75L	2e-24	66	83	0	AAT09735.1
61	377	349,952	Unknown protein	RV	CMTV 33R	1e-77	41	55	11	AFA44937.1
64	324	300,659	NTPase/ helicase	Other	SGIV 146L	5e-98	47	62	0	AAS18161.1
63[viii]	405	375,990	Unknown protein	RV	SGIV 147L	8e-17	25	42	17	AAS18162.1
65	216	199,789	Unknown protein	N/A	EHNV 37R	7e-28	37	59	3	ACO25227.1
66	79	73,398	Unknown protein	RV	SGIV 143L	4e-21	52	67	1	AAS18158.1
67	181	167,316	Unknown protein	RV	EHNV 35L	0.17	35	49	9	ACO25225.1
69	230	213,301	Unknown protein	Other	GIV 74R	1e-53	40	64	4	AAV91089.1
70	168	155,868	Putative myeloid cell leukemia protein	Other	EHNV 66R	9e-05	45	59	0	ACO25256.1
72[ix]	1,177	1,092,746	Unknown protein	Core	EHNV 77R	0.0	60	75	1	ACO25267.1
73	119	111,651	Unknown protein	RV	FV3 14R	6e-34	51	74	0	AAT09673.1
74	300	278,701	ATPase-like protein	Core	EHNV 92L	4e-166	82	90	0	ACO25282.1
78	154	142,889	Unknown protein	RV	CMTV 77L	3e-19	37	50	2	AFA44983.1
82	270	251,834	eIF2a-like protein	F/A/R	EHNV 61R	2e-24	31	45	11	ACO25251.1
85[x]	466	434,344	Putative ATPase-dependent protease	RV	ATV 26L	4e-116	50	65	2	AAP33203.1

Table 3.3 (continued): Predicted LMBV open reading frames (ORFs).

86[xi]	957	888,216	Putative tyrosine kinase	Core	RGV 29R	0.0	50	60	6	AFG73071.1
87	38	35,122	Unknown protein	RV	SGIV 59L	4e-04	61	79	0	AAS18074.1
88[xii]	1,023	952,079	Putative NTPase	Core	FV3 9L	0.0	67	79	0	AAT09668.1
94	356	330,967	Putative XPPG-RAD2-type nuclease	Core	STIV 100R	5e-161	60	76	1	ACF42318.1
95	160	149,126	Hypothetical protein-Clostridium tetani	Core	ESV 18R	3e-56	77	87	0	AFJ52301.1
96	267	246,002	Unknown protein	Other	SGIV 20L	2e-09	36	55	2	AAS18035.1
97	136	127,051	Unknown protein	Other	CMTV 63R	1e-55	63	78	1	AFA44969.1
99[xiii]	445	413,066	Helicase-like protein	Other	TFV 56L	8e-131	50	64	10	AAK55107.1
102	136	126,143	Unknown protein	N/A	SGIV 151L	4e-24	39	59	4	AAS18166.1
103	431	400,293	Serine-threonine protein kinase	Core	EHNV 48L	1e-104	47	63	2	ACO25238.1
105	79	73,565	Unknown protein	N/A	GIV 38R	1e-07	42	66	8	AAV91059.1
107	205	191,050	NIF/NLI interacting factor	Core	TFV 40R	8e-82	61	74	0	ABB92302.1
113[xiv]	1,307	1,215,485	DNA-dependent RNA polymerase II largest subunit	Core	ESV 11R	0.0	64	75	4	AFJ52294.1
114[xv]	1,094	1,016,503	DNA-dependent RNA polymerase II second largest subunit	Core	ATV 43R	0.0	71	80	3	AAP33221.1
117	340	316,480	Ribonuclease III	Core	FV3 80L	2e-130	62	76	1	AAT09740.1
121	91	84,433	Transcription elongation factor TFIIS	Core	STIV 88R	3e-24	56	67	2	ACF42306.1
122	147	136,251	Immediate early protein ICP-18	RV	SGIV 86R	2e-4149	49	64	5	AAS18101.1
123[xvi]	196	181,514	Unknown protein	N/A	ESV 81L	6e-25	35	52	4	AFJ52365.1
125	220	204,685	Cytosine DNA methyl-transferase	Other	TFV 89R	1e-93	67	84	0	AAL77813.1
126[xvii]	191	177,153	K2R protein from Cowpox	N/A	N/A	8e-06	40	51	1	CAD90750.1
131[xviii]	1,004	934,455	DNA polymerase	Core	STIV 63R	0.0	76	86	1	ACF42281.1

Table 3.3 (continued): Predicted LMBV open reading frames (ORFs).

132	72	67,559	Unknown protein	N/A	N/A	N/A	N/A	N/A	N/A	N/A
134	74	67,398	Unknown protein	N/A	N/A	N/A	N/A	N/A	N/A	N/A
136	382	354,406	Immediate early protein ICP-46	Core	EHNV 13L	7e-167	58	74	0	ACO25203.1
138	463	427,236	Major capsid protein	Core	ESV 22L	0.0	85	91	0	ACO90017.1
139[xix]	364	338,633	Unknown protein	Other	FV3 3R	1e-29	53	70	1	AAT09662.1
141	290	269,134	Unknown protein	N/A	ESV 2L	1e-79	46	61	4	AFJ52285.1
142	359	332,295	Putative myristylated membrane protein	Core	ESV 1L	1e-146	76	87	0	AFJ52284.1
145	50	46,562	Unknown protein	N/A	SGIV 138L	0.15	34	63	0	AAS18153.1
147	66	61,216	Unknown protein	F/A/R	EHNV 34L	4e-06	43	57	14	ACO25224.1
150	89	82,783	Unknown protein	N/A	N/A	N/A	N/A	N/A	N/A	N/A
153	200	186,314	Neuro-filament triplet H1-like protein	Other	STIV 35R	2e-24	51	72	4	ACF42253.1
155	62	57,470	Unknown protein	Other	FV3 33R	4e-12	64	79	0	AAT09692.1
156	99	92,117	Unknown protein	RV	EHNV 70R	1e-08	49	57	21	ACO25260.1
157	204	190,773	Collagen-like	Other	EHNV 39R	2e-07	76	90	0	ACO25229.1
159	407	378,594	Ribo-nucleotide reductase small subunit	Core	RGV 73L	0.0	79	86	0	AAS67856.1
163	253	234,896	Putative replication factor and/or DNA binding-packing	Core	EHNV 100R	1e-106	63	78	1	ACO25290.1
168	235	218,683	Putative integrase-like protein	N/A	CMTV 92R	9e-45	50	61	3	AFA44998.1
169	132	123,689	Unknown protein	N/A	SGIV 45L	2e-04	60	65	0	AAS18060.1
174	75	70,263	Unknown protein	N/A	STIV 77R	1e-21	59	68	0	ACF42295.1
177	91	84,785	CARD-like capsase	RV	ATV 40L	1e-15	41	64	0	ACB11347.1
196[xx]	532	494,940	Methyl-accepting chemotaxis sensory transducer	RV	STIV 18L	4e-178	52	73	6	AFG73061.1
198	274	254,496	Tumor necrosis factor receptor	N/A	SGIV 51L	1e-36	37	54	1	AAS18066.1

Table 3.3 (continued): Predicted LMBV open reading frames (ORFs).

224[xxi]	74	47,323	Unknown protein	Other	ESV 64L	3e-17	74	91	0	AFJ52 348.1
425	92	85,448	Unknown protein	N/A	SGIV 158L	0.001	36	49	16	AAS1 8173.1

[i] ORFs 6+115 aligned; values from blastx for combined sequence
[ii] ORF 16 +17 aligned; values from blastx for combined sequence
[iii] ORF 18 + 63 not aligned, correspond to Eaton FV3 23R, 24R; values from individual blastx results
[iv] ORFs 40+211 aligned; values from blastx for combined sequence
[v] ORFs 44+164 aligned; values from blastx for combined sequence
[vi] ORFs 48+49 aligned; values from blastx for combined sequence
[vii] ORFs 53+192 combined; values from blastx for combined sequences
[viii] ORF 18 + 63 not aligned, correspond to Eaton FV3 23R, 24R; values from individual blastx results
[ix] ORFs 72+241+171 aligned; values from blastx for combined sequence
[x] ORFs 85+181+116 aligned; values from blastx for combined sequence
[xi] ORFs 86+80+25+152 aligned; values from blastx for combined sequence
[xii] ORFs 88+109 aligned; values from blastx for combined sequence
[xiii] ORFs 99 + 23 aligned; values from blastx for combined sequence
[xiv] ORFs 113+26 aligned; values from blastx for combined sequence
[xv] ORFs 114+173+92+30 aligned; values from blastx for combined sequence
[xvi] ORFs 123+124 came back as ESV 81L; used ORF123 because of lower expect value; however, this ORF and the ESV 81L ORF may be two instances of cellular contamination.
[xvii] Blast to cowpox gene
[xviii] ORFs 131+54 aligned; values from blastx for combined sequence
[xix] ORFs 139 + 75 aligned; values from blastx for combined sequence
[xx] ORF 196+11 aligned; values from blastx for combined sequence
[xxi] ORFs 252+234 aligned; values from blastx for combined sequence

Table 3.3 (continued): Predicted LMBV open reading frames (ORFs).

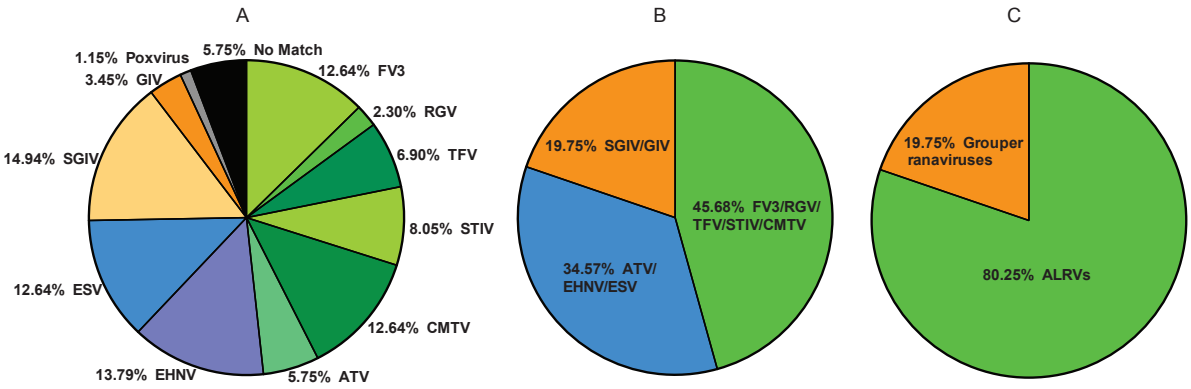

Figure 3.1: Best BLASTX matches for predicted LMBV ORFs illustrate a higher similarity to ALRVs. The pie chart in panel A illustrates the percentage of predicted LMBV ORFs that best matched genes in complete Ranavirus genomes, complete poxvirus genomes, and those with no identified match. Panel B shows the percentage of predicted LMBV ORFs that best matched the grouper ranaviruses and two separate clades of ALRVs (EHNV, ATV, ESV) and (FV3, STIV, RGV, TFV, CMTV). Panel C displays the percentage of predicted LMBV ORFs that best matched the grouper ranaviruses and all ALRVs.

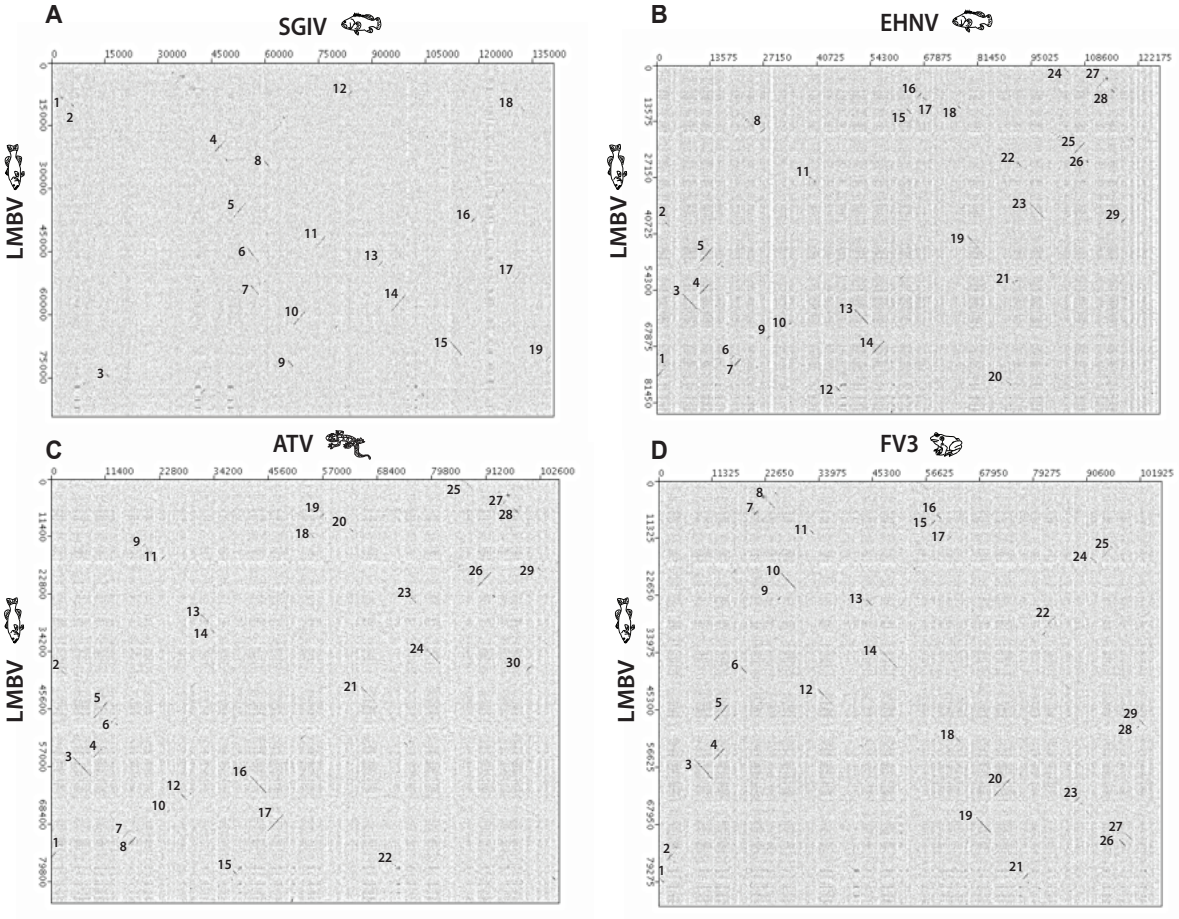

Figure 3.2: Dot plot analyses of the LMBV sequence compared to the complete genomic sequences from grouper and ALRVs indicate limited collinearity. Dot plots were generated comparing the LMBV sequence to SGIV (A), EHNV (B), ATV (C), and FV3 (D). Segments of collinearity, indicated by a straight line, were sequentially numbered along the non-LMBV genomic sequences.

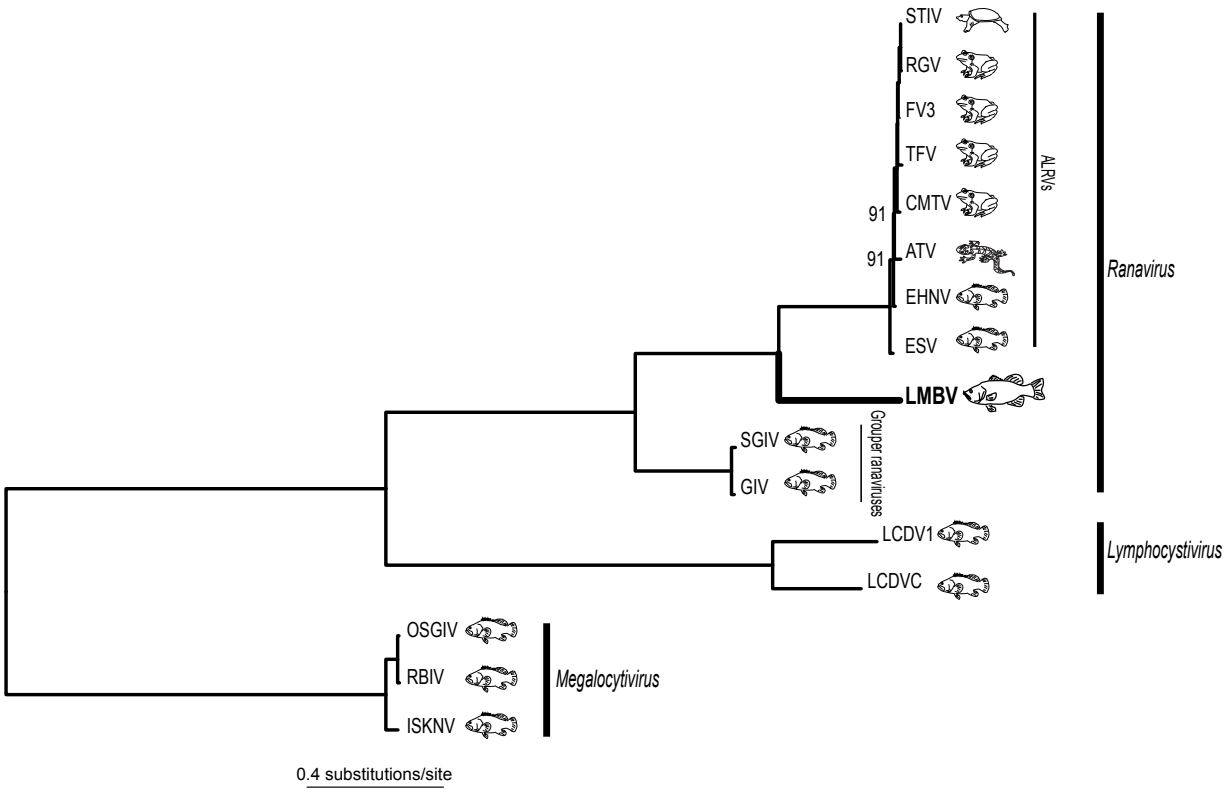

Figure 3.3: Phylogenetic relationships among vertebrate Iridoviridae based on the 26 core Iridoviridae genes indicate a closer LMBV-ALRV relationship. A phylogenetic tree of viral species/strains isolated from fish, amphibians, and reptiles based on a maximum likelihood analysis of a concatenated data set of 26 Core *Iridoviridae* genes. The LMBV branch is highlighted in bold. Bootstrap values < 100 are shown. Note that each fish, amphibian, and reptile symbol denotes multiple species as indicated in Table 3.2.

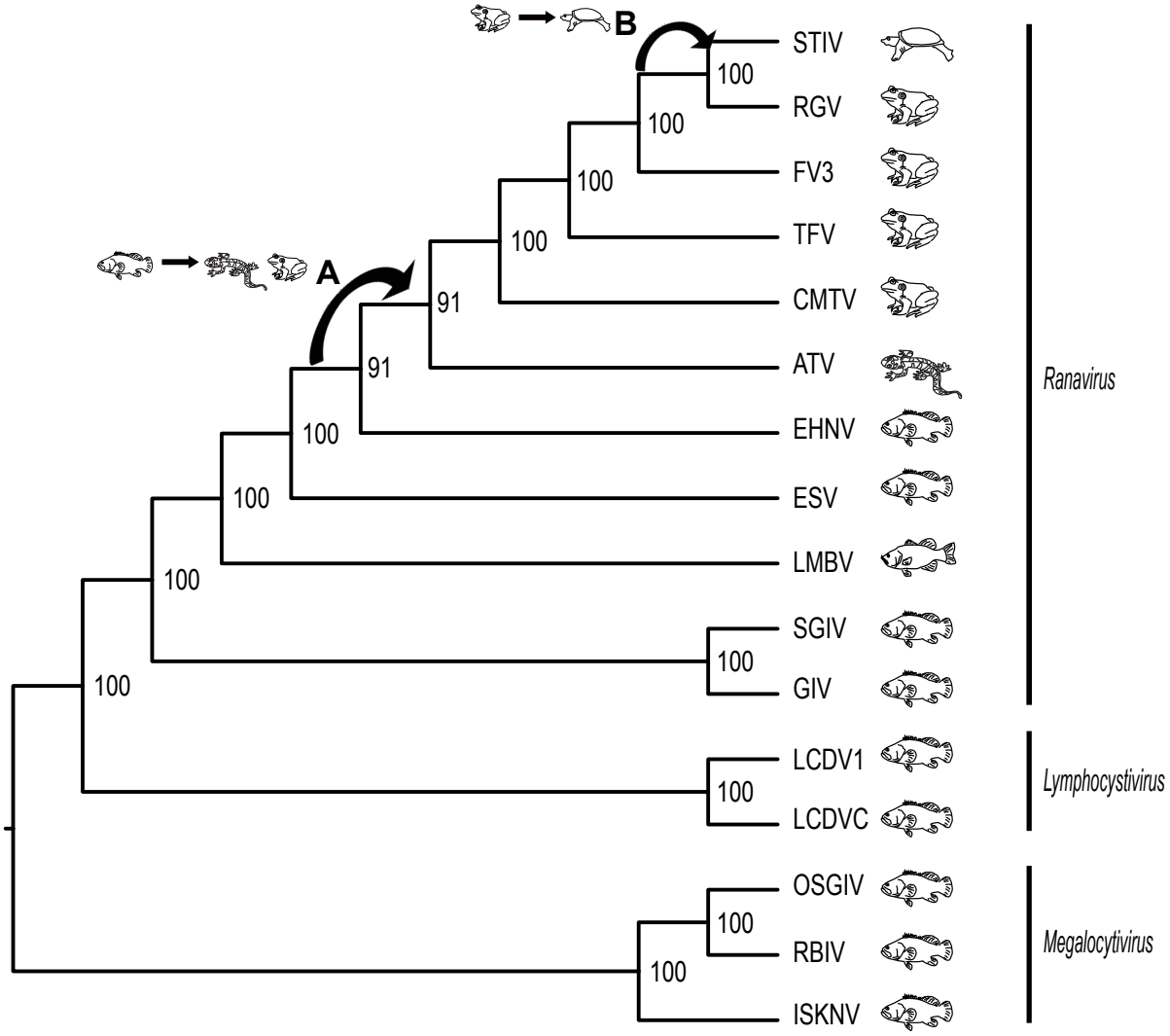

Figure 3.4: Phylogenetic relationships among vertebrate *Iridoviridae* suggests two major host switches within *Ranavirus*. A phylogenetic tree of viral species/strains isolated from fish, amphibians, and reptiles based on a maximum likelihood analysis of a concatenated data set of 26 Core *Iridoviridae* genes. Arrow A shows a clear transition from fish hosts to salamander/frog hosts and arrow B indicates a transition from frog hosts to turtle hosts. Bootstrap values are shown. Note that each fish, amphibian, and reptile symbol denotes multiple species as indicated in Table 3.2.

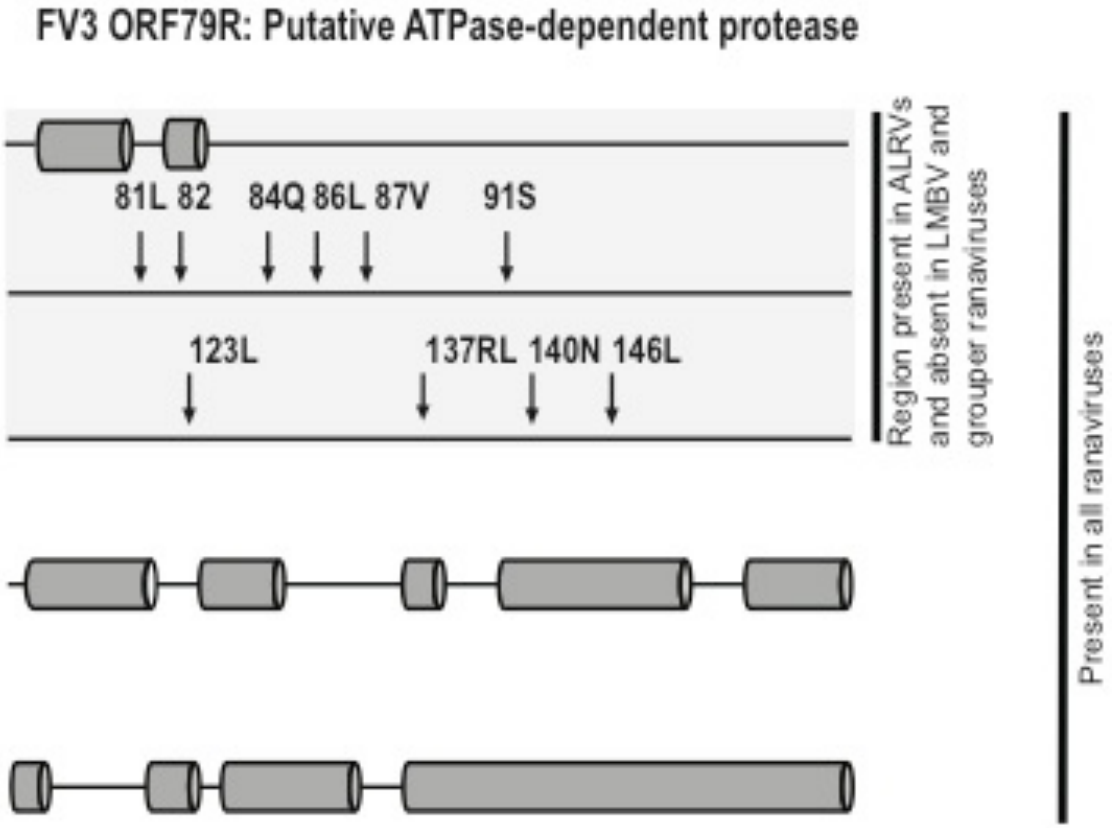

Figure S3.1: Secondary protein structure for FV3 ORF79R/Putative ATPase-dependent protease shows the region with positively selected sites is only present in ALRVs. Predicted secondary structures for FV3 ORF 79R/Putative ATPase-dependent protease is shown. α-helices (cylinders), β-strands (horizontal arrows), and unstructured regions (solid lines) are illustrated. Vertical arrows indicated sites under positive selection (Abrams *et al*. 2013). The region of the protein absent in LMBV and grouper ranaviruses is highlighted with a gray box. This figure is adapted from Abrams *et al*. (2013) with permission.

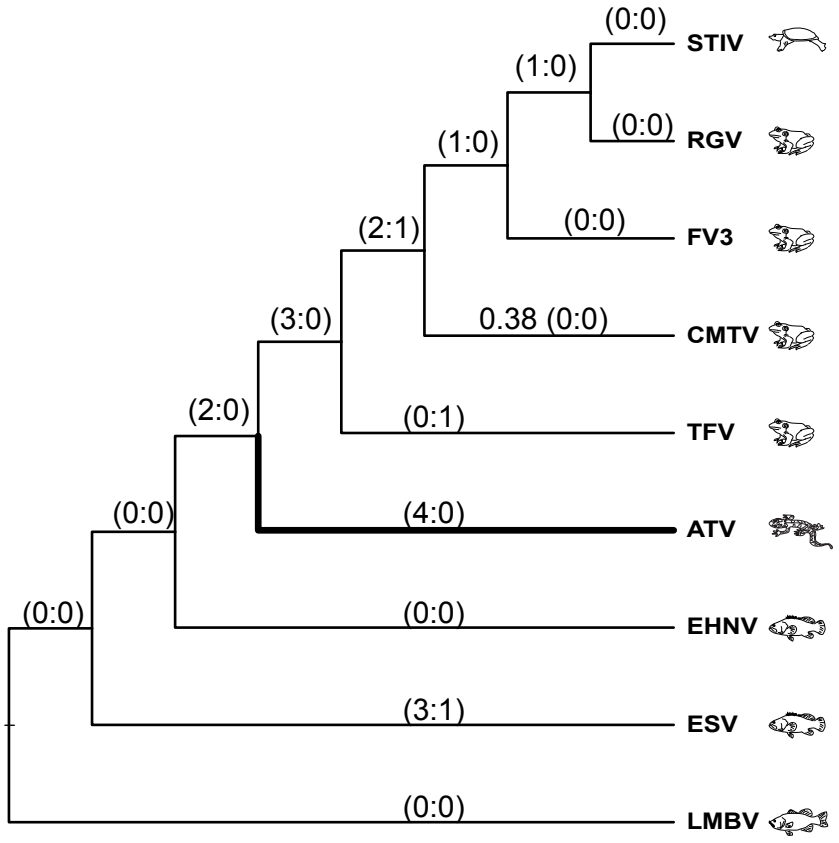

Figure S3.2: Elevated *dN*/*dS* identified along the ATV branch for the F/A/R gene ORF 71R. A phylogeny for FV3 ORF 71R with the addition of the predicted LMBV sequence is presented. Branch values of *dN*/*dS* are shown, along with the estimated numbers of non-synonymous and synonymous (*N*:*S*) mutations that are predicted to have occurred along each branch (in parentheses). Branches with incalculable *dN*/*dS* values (due to *S*=0) are highlighted in bold if the ratio of N:S is greater than or equal to 4:0; therefore the ATV branch is highlighted in bold due to an estimated dN/dS value greater than 1. Note variation in the topology of FV3 ORF 71R due to a slightly different gene tree.

References

Abrams, A.J., Cannatella, D.C, Hillis, D.M, & Sawyer, S.L 2013. Recent host-shifts in ranaviruses: signatures of positive selection in the viral genome. *J Gen Virol* 94:2082-2093.

Aitcheson, S.J., Arnett, J., Murray, K.R., & Zhang, J. 2000. Removal of aquaculture therapeutants by carbon adsorption: 1. Equilibrium adsorption behaviour of single components. *Aquaculture* 183:269-284.

Anisimova M., Bielawski J., & Yang Z. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol Biol and Evol*. 18:1585-1592.

Babe, L.M., & Craik, C.S. 1997. Viral proteases: evolution of diverse structural motifs to optimize function. *Cell* 91:427-430.

Bandin I. & Dopazo C. 2011. Host range, host specificity and hypothesized host shift events among viruses of lower vertebrates. *Vet Res*. 42:1-15.

Bayley, A.E., Hill, B.J., & Feist, S.W. 2013. Susceptibility of the European common frog *Rana temporaria* to a panel of ranavirus isolates from fish and amphibian hosts. *Dis Aquat Org*. 103:171-183.

Becker, J.A., Tweedie, A., Gilligan, D., Asmus, M., & Whittington, R.J. 2013. Experimental infection of Australian freshwater fish with Epizootic Haematopoietic Necrosis Virus (EHNV). *J Aquat Anim Health* 25:66-76.

Bhatt S., Lam T., Lycett S., Brown A., Bowden T., Holmes E. C., Guan Y., Wood J., Brown I., Kellum P., *et al*. 2013. The evolutionary dynamics of influenza A virus adaptation to mammalian hosts. *Phil Trans R Soc B*. 368, 20120382.

Blazer, V.S., Iwanowicz, L.R., Starliper, C.E., Iwanowicz, D.D., Barbash, P., Hedrick, J.D., Reeserd, S.J., Mullicane, J.E., Zauggf, S.D., Burkhardtf, M.R., & Kelble, J. 2010. Mortality of centrarchid fishes in the Potomac drainage: survey results and overview of potential contributing factors. *J Aquat Anim Health* 22:190-218.

Bly, J.E. Quiniou, S.M. & Clem, L.W. 1997. Environmental effects on fish immune mechanisms. *Dev Biol Stand*. 90:33-43.

Brodie, R., Roper, R.L., & Upton, C. 2004. JDotter: a Java interface to multiple dotplots generated by dotter. *Bioinformatics* 20:279-281.

Brunner, J.L., Schock, D.M., Davidson, E.W., & Collins, J.P. 2004. Intraspecific reservoirs: complex life history and the persistence of a lethal *Ranavirus*. *Ecology* 85:560-566.

Cantino, P.D. and de Queiroz, K. 2010. International Code of Phylogenetic Nomenclature, version 4c.

Carey, C., Cohen, N., & Rollins-Smith, L. 1999. Amphibian declines: an immunological perspective. *Dev Comp Immunol*. 23:459-472.

Carmichael, G.J. & Tomasso, J.R. 1983. Use of formalin to separate tadpoles from largemouth bass fingerlings after harvesting. *Prog Fish-Cult*. 45:105-106.

Chaijan, M., Benjakul, S., Visessanguan, W., Lee, S., & Faustman, C. 2007. The effect of freezing and aldehydes on the interaction between fish myoglobin and myofibrillar proteins. *J Agric Food Chem*. 55:4562-4568.

Chen, G., Ward, B.M., Kwang, H.Y., Chinchar, V.G., & Robert, J. 2011. Improved knockout methodology reveals that frog virus 3 mutants lacking either the 18K immediate-early gene or the truncated vIF-2α gene are defective for replication and growth in vivo. *J Virol*. 85:11131-11138.

Chinchar V.G. 2002. Ranaviruses (family *Iridoviridae*): emerging cold-blooded killers. *Arch Virol*. 147:447-470.

Chinchar, V.G., Hyatt, A., Miyazaki, T., & Williams, T. 2009. Family Iridoviridae: poor viral relations no longer. In *Lesser Known Large dsDNA Viruses* (pp. 123-170). Springer Berlin Heidelberg.

Choe H., Jemielity S., Abraham J., Radoshitzky S., & Farzan M. 2011. Transferrin receptor 1 in the zoonosis and pathogenesis of New World hemorrhagic fever arenaviruses. *Curr Opin Microbiol*. 14:476-482.

Collins J. & Storfer A. 2003. Global amphibian declines: sorting the hypotheses. *Divers Distrib*. 9:89–98

Corse, W. & Metter, E. 1980. Economics, adult feeding and larval growth of *Rana catesbeiana* on a fish hatchery. *J Herpetol*. 14:231-238.

Cullen, B.R. & Owens, L. 2002. Experimental challenge and clinical cases of Bohle iridovirus (BIV) in native Australian anurans. *Dis Aquat Org*. 49:83-92.

Cooke, S.J., Schreer, J.F., Wahl, D.H., & Philipp, D.P. 2002. Physiological impacts of catch-and-release angling practices on largemouth bass and smallmouth bass. In *American Fisheries Society Symposium* (pp. 489-512). American Fisheries Society.

Culp, S.J. & Beland, F.A. 1996. Malachite green: a toxicological review. *International Journal of Toxicology* 15:219-238.

Daszak P., Berger L., Cunningham A., Hyatt A., Green D., & Speare R. 1999. Emerging infectious diseases and amphibian population declines. *Emerg Infect Dis*. 5:735-748.

Daszak, P., Cunningham, A.A. & Hyatt, A.D. 2000. Emerging infectious diseases of wildlife—threats to biodiversity and human health. *Science* 287:443-449.

Delport W., Poon A., Frost S., & Pond S. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26:2455-2457.

Delcher, A.L., Harmon, D., Kasif, S., White, O., & Salzberg, S.L. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res*. 27:4636-4641.

Demogines A., Farzan M., & Sawyer S. 2012a. Evidence for ACE2-utilizing coronaviruses (CoVs) related to severe acute respiratory syndrome CoV in bats. *J Virol*. 86:6350-6353.

Demogines A., Truong K., & Sawyer S. 2012b. Species-specific features of DARC, the primate receptor for Plasmodium vivax and Plasmodium knowlesi. *Mol Biol and Evol*. 29:445-449.

Demogines A., Abraham J., Choe H., Farzan M., & Sawyer S. 2013. Dual host-virus arms races shape an essential housekeeping protein. *PLoS Biol* 11: e1001571. doi:10.1371/journal.pbio.1001571.

Deng, G., Li, S., Xie, J., Bai, J., Chen, K., Ma, D., Jiang, X., Lao, H., & Yu, L. 2011. Characterization of a ranavirus isolated from cultured largemouth bass (*Micropterus salmoides*) in China. *Aquaculture* 312:198-204.

Do J.W., Moon C.H., Kim HJ.., Ko, M.S., Kim, S.B., Son, J.H., Kim, J.S., An, E.J., Kim, M.K. & others authors. 2004. Complete genomic DNA sequence of rock bream iridovirus. *Virology* 325:351-363.

Docherty, D.E., Meteyer, C.U., Wang, J., Mao, J., Case, S.T., & Chinchar, V.G. 2003. Diagnostic and molecular evaluation of three iridovirus-associated salamander mortality events. *J Wildl Dis*. 39:556-566.

Eaton H.E., Metcalf J., Penny E., Tcherepanove V., Upton C., & Brunetti C.R. 2007. Comparative genomic analysis of the family Iridoviridae: re-annotating and defining the core set of iridovirus genes. *Virology J*. 4:1-17.

Elde N., Child S., Eickbush M., Kitzman J., Rogers K., Shendure J., Geballe A., & Malik H. 2012. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. *Cell* 150:831-841.

Ellsaesser, C.F. & Clem, L.W. 1986. Haematological and immunological changes in channel catfish stressed by handling and transport. *J Fish Biol*. 28:511-521.

Endo T., Ikeo K., & Gojobori T. 1996. Large-scale search for genes on which positive selection may operate. *Mol Biol and Evol*. 13:685-690.

Forson, D.D & Storfer, A. 2006. Atrazine increases *Ranavirus* susceptibility in the tiger salamander, *Ambystoma tigrinum*. *Ecol Appl*. 16:2325-2332.

Gifford R. 2011. Viral evolution in deep time: lentiviruses and mammals. *Trends in Genet*. 28:89-100.

Goldberg, T.L., Coleman, D.A., Grant, E.C., Inendino, K.R., & Philipp, D.P. 2003. Strain variation in an emerging iridovirus of warm-water fishes. *J Virol*. 77:8812-8818.

Goldberg, T.L., Grant, E.C., Inendino, K.R., Kassler, T.W., Claussen, J.E., & Philipp, D.P. 2005. Increased infectious disease susceptibility resulting from outbreeding depression. *Conserv Biol*. 19:455-462.

Gray, M.J, Miller, D.L., Schmutzer, A.C. & Baldwin, C.A. 2007. Frog virus 3 prevalence in tadpole populations inhabiting cattle-access and non-access wetlands in Tennessee, USA. *Dis Aquat Organ*. 77:97-103.

Grayfer L., Andino F., Chen G., Chinchar G., & Robert J. 2012. Immune evasion strategies of ranaviruses and innate immune responses to these emerging pathogens. *Viruses* 4:1075-1092.

Green, D.E., Converse, K.A. & Schrader, A.K. 2002. Epizootiology of sixty-four amphibian morbidity and mortality events in the USA, 1996-2001. *Ann N Y Acad Sci*. 969:323-339.

Grizzle, J.M., Altinok, I., Fraser, W.A., & Francis-Floyd, R. 2002. First isolation of largemouth bass virus. *Dis. Aquat. Organ*. 50: 233–235.

Gobbo, F., Cappellozza, E., Pastore, M.R. & Bovo, G. 2010. Susceptibility of black bullhead *Ameirus melas* to a panel of Ranavirus isolates. *Dis Aquat Organ*. 90:167-174.

Goldman N. & Yang Z. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol Biol and Evol*. *11*:725-736.

Hanson, L.A., Petrie-Hanson, L., Meals, K.O., Chinchar, V.G., & Rudis, M. 2001. Persistence of largemouth bass virus infection in a northern Mississippi reservoir after a die-off. *J Aquat Anim Health* 13:27-34.

Hanselmann, R., Rodrıguez, A., Lampo, M., Fajardo-Ramos, L., Alonso Aguirre, A., Kilpatrick, A., Rodriguez, J. & Daszak, P. 2004. Presence of an emerging pathogen of amphibians in introduced bullfrogs *Rana catesbeiana* in Venezuela. *Biol Cons*. 120:115-119.

Harris, J. & Bird, D.J. 2000. Modulation of the fish immune system by hormones. *Vet Immunol Immunop*. 77:163-176.

Harrison R. & Bonning B. 2004. Application of maximum-likelihood models to selection pressure analysis of group 1 nucleopolyhedrovirus genes. *J Virol*. 85:197-210.

He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z., & Chan S.M. 2001. Complete genome analysis of mandarin fish infectious spleen and kidney necrosis iridovirus. *Virology* 291:126-139.

He J.G., Lu L., Deng M., He H.H., Weng S.P., Wang X.H., Zhou S.Y., Long Q.X., Wang X.Z., & Chan S.M. 2002. Sequence analysis of the complete genome of an iridovirus isolated from the tiger frog. *Virology* 292:185-197.

Herfst S., Schrauwen E., Linster M., Chutinimitkul, S., de Wit, E., Munster, V., Sorrell, E., Bestebroer, T., Burke, D. & other authors. 2012. Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* 336:1534-1541.

Hodgson, J.R. & Hansen, E.M. 2005. Terrestrial prey items in the diet of largemouth bass, *Micropterus salmoides*, in a small north temperate lake. *Journal of Freshwater Ecology* 20:793-794.

Hoelzer K., Shackelton L., Parrish C., & Holmes E. 2008. Phylogenetic analysis reveals the emergence, evolution and dispersal of carnivore parvoviruses. *J Gen. Virol.* 89:2280-2289.

Hoverman, J.T., Gray, M.J., Haislip, N.A., & Miller, D.L. 2011. Phylogeny, life history, and ecology contribute to differences in amphibian susceptibility to ranaviruses. *EcoHealth* 8:301-319.

Huang Y., Huang X., Liu H., Gong J., Ouyang Z., Cui H., Cao J., Zhao Y., Wang X., Jiang Y., & Qin Q. 2009. Complete sequence determination of a novel reptile iridovirus isolated from soft-shelled turtle and evolutionary analysis of *Iridoviridae*. *BMC Genomics* 10:224.

Hyatt, A.D., Gould, A.R., Zupanovic, Z., Cunningham, A.A., Hengstberger, S., Whittington, R.J., Kattenbelt, J., & Coupar, B.E. 2000. Comparative studies of piscine and amphibian iridoviruses. *Arch Virol*. 145:301-331.

Iwanowicz, L., Densmore, C., Hahn, C., McAllister, P., & Odenkirk, J. 2013. Identification of largemouth bass virus in the introduced northern snakehead inhabiting the Chesapeake Bay watershed. *J Aquat Anim Health* 25:191-196.

Jancovich, J.K., Davidson, E.W., Seiler, A., Jacobs, B.L., & Collins, J.P. 2001. Transmission of the *Ambystoma tigrinum* virus to alternative hosts. *Dis Aquat Organ*. 46:159-163.

Jancovich J.K., Mao J., Chinchar V.G., Wyatt C., Case S.T., Kumar S., Valente G., Subramanian S., Davidson E.W., Collins J.P., & Jacobs, B.L. 2003. Genomic sequence of a *Ranavirus* (family *Iridoviridae*) associated with salamander mortalities in North America. *Virology* 316:90-103.

Jancovich J.K., Davidson E.W., Parameswaran N., Mao J., Chinchar V.G., Collins J.P., Jacobs B.L., & Storfer A. 2004. Evidence for emergence of an amphibian iridoviral disease because of human-enhanced spread. *Mol Ecol*. 14:213-224.

Jancovich J.K., Bremont M., Touchman J.W., & Jacobs B.L. 2010. Evidence for multiple recent host species shifts among the ranaviruses (Family Iridoviridae). *J Virol*. 84:2636–2647.

Jensen, B.B., Ersboll, A.K., & Ariel, E. 2009. Susceptibility of pike *Esox lucius* to a panel of *Ranavirus* isolates. *Dis Aquat Org*. 83:169-179.

Kaelber J., Demogines A., Harbison C., Allison A., Goodman L., Ortega A., Sawyer L., & Parrish C. 2012. Evolutionary reconstructions of the transferrin receptor of Caniforms supports canine parvovirus being a re-emerged and not a novel pathogen in dogs. *PLoS Pathog*. 8:e1002666.

Kelley L. & Sternberg M. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nature Proto*. 4:363-371.

Kerby, J.L. & Storfer, A. 2009. Combined effects of atrazine and chloropyrifos on susceptibility of the tiger salamander to *Ambystoma tigrinum* virus. *EcoHealth* 6:91-98.

Kiesecker, J.M., Blaustein, A.R., & Miller, C.L. 2001. Transfer of a pathogen from fish to amphibians. *Conserv Biol*. 15:1064-1070.

King, A.M., Adams, M.J., Carstens, E.B. & Lefkowitz, E.J. 2012. Virus taxonomy: classification and nomenclature of viruses: ninth report of the International Committee on Taxonomy of Viruses.

Kloskowski, J. 2010. Fish farms as amphibian habitats: factors affecting amphibian species richness and community structure at carp ponds in Poland. *Environ Conserv*. 37:187-194.

Kodama, H., Matsuoka, Y., Tanaka, Y., Liu, Y., Iwasaki, T., & Watarai, S. 2004. Changes of C- reactive protein levels in rainbow trout (*Oncorhynchus mykiss*) sera after exposure to anti-ectoparasitic chemicals used in aquaculture. *Fish Shellfish Immunol*. 16:589-597.

Kosakovsky Pond S., Posada D., Gravenor M., Woelk C., & Frost S. 2006. GARD: a genetic algorithm for recombination detection. *Bioinformatics* 22:3096-3098.

Kosiol C., Vinar T., da Fonseca R., Hubisz M., Bustamante C., Nielsen R., & Siepel A. 2008. Patterns of positive selection in six mammalian genomes. *PLOS Genetics* 4:e1000144.

Langdon, J.S., Humphrey, J.D., Williams, L.M., Hyatt, A.D., & Westbury, H.A. 1986. First virus isolation from Australian fish: an iridovirus-like pathogen from redfin

perch, *Perca fluviatilis*. L. *J Fish Dis*. 9:263–268.

Larkin M.A., Blackshields G., Brown N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A. and other authors. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947-2948.

Lei X.Y., Ou T., Zhu R.L., & Zhang Q.Y. 2012. Sequencing and analysis of the complete genome of Rana grylio virus (RGV). *Arch Virol*. 157:1559-1564.

Lesbarrères D., Balseiro A., Brunner J., Chinchar, G., Duffus, A., Kerby, J., Miller, D., Robert, J., Schock, D. and other authors. 2012. Ranavirus: past, present and future. *Biol Let*. 8:481-483.

Lu L., Zhou S.Y., Chen C., Weng S.P., Chan S.M., & He J.G. 2005. Complete genome sequence analysis of an iridovirus isolated from the orange-spotted grouper, *Epinephelus coioides*. *Virology* 339:81-100.

Ma, T.H. & Harris, M.M. 1988. Review of the genotoxicity of formaldehyde. *Mutat Res-Rev Genet*. 196:37-59.

Maceina, M.J. & Grizzle, J.M. 2006. The relation of largemouth bass virus to largemouth bass population metrics in five Alabama reservoirs. *Trans Am Fish Soc*. 135:545-555.

Mao, J., Hedrick, R.P., & Chinchar, V.G.. 1997. Molecular characterization, sequence analysis, and taxonomic position of newly isolated fish iridoviruses. *Virology* 229:212-220.

Mao J., Green D.E., Fellers G., & Chinchar, V.G. 1999a. Molecular characterization of Iridoviruses isolated from sympatric amphibians and fish. *Virus Res*. 63:45–52.

Mao, J., Wang, J., Chinchar, G.D., & Chinchar, V.G. 1999b. Molecular characterization of a ranavirus isolated from largemouth bass *Micropterus salmoides*. *Dis Aquat Org*. 37:107-114.

Mavian C., López-Bueno A., Balseiro A., Casais R., Alcamí A., & Alejo A. 2012a. The genome sequence of the emerging common midwife toad virus identifies an evolutionary intermediate within ranaviruses. *J Virol*. 86:3617-3625.

Mavian C., López-Bueno A., Fernandez Somalo M.P., Alcamí A., & Alejo A. 2012b. Complete genome sequence of the european sheatfish sirus. *J Virol*. 86:6365-6366.

McLysaght A., Baldi P., & Gaut B. 2003. Extensive gene gain associated with adaptive evolution of poxviruses. *PNAS* 100:15655-15660.

Miller, D.L., Rajeev, S., Gray, M.J., & Baldwin, C.A. 2007. Frog virus 3 infection, cultured American bullfrogs. *Emerg Infect Dis*. 13:342.

Miller D., Gray M., & Storfer A. 2011. Ecopathology of ranaviruses infecting amphibians. *Viruses* 3:2351-2373.

Meinelt, T., Playle, R., Schreckenbach, K., & Pietrock, M. (2001). The toxicity of the antiparasitic mixture FMC is changed by humic substances and calcium. *Aquaculture Res*. 32:405-410.

Moody, N.J.G. & Owens, L. 1994. Experimental demonstration of the pathogenicity of a frog virus, Bohle iridovirus, for a fish species, barramundi *Lates calcarifer*. *Dis Aquat Organ*. 18:95-102.

Murray, A.G. & Peeler, E.J. 2005. A framework for understanding the potential for emerging diseases in aquaculture. *Prev Vet Med*. 67:223-235.

Neal, J.W., Eggleton, M.A., & Goodwin, A.E. 2009. The effects of largemouth bass virus on a quality largemouth bass population in Arkansas. *J Wildl Dis*. 45:766-771.

Nielsen R. & Yang, Z. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148:929–936.

Nilsson J., Grahn M., & Wright A. 2011. Proteome-wide evidence for enhanced positive Darwinian selection within intrinsically disordered regions in proteins. *Genome Biol*. 12:R65.

Ohlemeyer, S., Holopainen, R., Tapiovaara, H., Bergmann, S.M., & Schütze, H. 2011. Major capsid protein gene sequence analysis of the Santee-Cooper ranaviruses DFV, GV6, and LMBV. *Dis Aquat Org*. 96:195-207.

Parrish C., Holmes E., Morens D., Park E., Burke D., Calisher C., Laughlin C., Saif L., & Daszak, P. 2008. Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol Mol Biol R*. 72:457-470.

Peeler, E.J., Oidtmann, B.C., Midtlyng, P.J., Miossec, L., & Gozlan, R.E. 2011. Non-native aquatic animal introductions have driven disease emergence in Europe. *Biol Invasions* 13:1291-1303.

Picco A.M. & Collins J.P. 2008. Amphibian commerce as a likely source of pathogen pollution. *Conserv Biol*. 22:1582-1589.

Picco, A.M., Karam, A.P., & Collins, J.P. 2010. Pathogen host switching in commercial trade with management recommendations. *EcoHealth* 7:252-256.

Plumb, J.A., Grizzle, J.M., Young, H.E., Noyes, A.D., & Lamprecht, S. 1996. An Iridovirus isolated from wild largemouth bass. *J Aquat Anim Health* 8:265–270.

Plumb, J.A., Noyes, A.D., Graziano, S., Wang, J., Mao, J., & Chinchar, V.G. 1999. Isolation and identification of viruses from adult largemouth bass during a 1997–1998 survey in the southeastern United States. *J Aquat Anim Health* 11:391-399.

Plumb, J.A. & Zilberg, D. 1999. The lethal dose of largemouth bass virus in juvenile largemouth bass and the comparative susceptibility of striped bass. *J Aquat Anim Health* 11:246-252.

Rachowicz L., Hero J., Alford R., Taylor J., Morgan J., Vredenburg V., Collins J., & Briggs C. 2005. The novel and endemic pathogen hypotheses: competing explanations for the origin of emerging infectious diseases of wildlife. *Conserv Biol*. 19:1441-1448.

Sabeti P., Schaffner S., Fry B., Lohmueller J., Varilly P., Shamovsky O., Palma A., Mikkelsen T., Altshuler D., & Lander E. 2006. Positive natural selection in the human lineage. *Science* 312:1614-1620.

Sawyer S. & Elde N. 2012. A cross-species view on viruses. *Curr Opin Virol*. 2:561–568.

Schloegel, L.M., Picco, A.M., Kilpatrick, A.M., Davies, A.J., Hyatt, A.D., & Daszak, P. 2009. Magnitude of the US trade in amphibians and presence of *Batrachochytrium dendrobatidis* and ranavirus infection in imported North American bullfrogs (*Rana catesbeiana*). *Biol Cons*. 142:1420-1426.

Schock, D.M., Bollinger, T.K., Chinchar, V.G., Jancovich, J.K., & Collins, J.P. 2008. Experimental evidence that amphibian ranaviruses are multi-host pathogens. *Copeia* 1:131-141.

Schramm, H.L., Walters, A.R., Grizzle, J.M., Beck, B.H., Hanson, L.A., & Rees, S.B. 2006. Effects of live well conditions on mortality and largemouth bass virus prevalence in largemouth bass caught during summer tournaments. *North Am J Fish Mana*. 26:812-825.

Shackelton, L., Parrish, C., Truyen, U., & Holmes, E. 2005. High rate of viral evolution associated with the emergence of carnivore parvovirus. *PNAS* 102:379-384.

Shchelkunov, S.N. 2003. Immunomodulatory proteins of orthopoxviruses. *Mol Biol*. 37:37-48.

Silveira-Coffigny, R., Prieto-Trujillo, A., & Ascencio-Valle, F. 2004. Effects of different stressors in haematological variables in cultured *Oreochromis aureus*. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 139:245-250.

Slobedman, B., Barry, P., Spencer, J., Avdic, S., & Abendroth, A. 2009. Virus-encoded homologs of cellular interleukin-10 and their control of host immune function. *J Virol*. 83:9618-9629.

Song W.J., Qin Q.W., Qiu J., Huang C.H., Wang F., & Hew, C.L. 2004. Functional genomics analysis of Singapore grouper iridovirus: complete sequence determination and proteomic analysis. *J Virol*. 78:12576-12590.

Southard, G.M., Fries, L.T., & Terre, D.R. 2009. Largemouth bass virus in Texas: distribution and management issues. *J Aquat Anim Health* 21:36-42.

Srivastava, S., Sinha, R., & Roy, D. 2004. Toxicological effects of malachite green. *Aquat Toxicol*., 66:319-329.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688-2690.

Stuart S., Chanson J., Cox N., Young B., Rodrigues A., Fischman D., & Waller R. 2004. Status and trends of amphibian declines and extinctions worldwide. *Science* 306:1783-1786.

Sudova, E., Machova, J., Svobodova, Z., & Vesely, T. 2007. Negative effects of malachite green and possibilities of its replacement in the treatment of fish eggs and fish: a review. *Veterinarni Medicina* 52:527-539

Tan W.G., Barkman T.J., Chinchar V.G., & Essani K. 2004. Comparative genomic analyses of frog virus 3, type species of the genus *Ranavirus* (family *Iridoviridae*). *Virology* 323:70-84.

Tidona CA.. & Darai G. 1997. The complete DNA sequence of lymphocystis disease virus. *Virology* 230:207-216.

Tsai C.T., Ting J.W., Wu M.H., Wu M.F., Guo I., & Chang C.Y. 2005. Complete genomic sequence of the grouper iridovirus and comparison of genomic organization with those of other iridoviruses. *J Virol*. 79:2010-2023.

Wain L., Bailes E., Bibollet-Ruche F., Decker, J., Keele, B., Heuverswyn, F., Li, Y., Takehisa, J. Ngole, E. and other authors. 2007. Adaptation of HIV-1 to its human host. *Mol Biol and Evol*. 24:1853-1860.

Whittington, R.J., Becker, J.A., & Dennis, M.M. 2010. Iridovirus infection in finfish – critical review with emphasis on ranaviruses. *J Fish Dis*.^33:95-122.

Williams T., Barbosa-Solomieu V., & Chinchar, V.G. 2005. A decade of advances in iridovirus research. *Adv Virus Res*. 65:173-248.

Wolffe, E., Vijaya, S., & Moss, B. 1995. A myristylated membrane protein encoded by the vaccinia virus L1R open reading frame is the target of potent neutralizing monoclonal antibodies. *Virology* 211:53-63.

Woodland, J.E., Brunner, C.J., Noyes, A.D., & Grizzle, J.M. 2002. Experimental oral transmission of largemouth bass virus. *J Fish Dis*. 25:669-672.

Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol and Evol*. 15:568-573.

Yang, Z., Swanson W., & Vacquier V. 2000. Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. *Mol Biol and Evol*. 17:1446-455.

Yang Z. 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Mol Biol and Evol*. 24:1586-1591.

Yildiz, H.Y. & Pulatsü, S. 1999. Evaluation of the secondary stress response in healthy Nile tilapia (*Oreochromis niloticus L.*) after treatment with a mixture of formalin, malachite green and methylene blue. *Aquacult Res*. 30: 379-383.

Yonar, M.E., & Yonar, S.M. 2010. Changes in selected immunological parameters and antioxidant status of rainbow trout exposed to malachite green (*Oncorhynchus mykiss*, Walbaum , 1792). *Pestic Biochem Physiol*. 97:19-23.

Zerbino, D.R. & Birney, E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 18:821-829.

Zhang Q.Y., Xiao F., Xie J., Li Z.Q., & Gui J.F. 2004. Complete genome sequence of lymphocystis disease virus isolated from China. *J Virol*. 78:6982-6994.