The Distribution, Prevalence, and Transmission of Ranaviruses and their Effect on Amphibian Communities of Vermont

Lauren V. Ash

Second Year Ph.D. Student

A proposal submitted to the committee members of Lauren V. Ash in partial fulfillment of the requirements for the Ph.D. qualification exam

Ph.D. Advisors Dr. Nicholas J. Gotelli

Dr. C. Brandon Ogbunugafor

Committee Members Dr. Samuel V. Scarpino, Committee Chair

Dr. Joseph J. Schall

Biology Department

University of Vermont

Burlington, VT

May 2017

**Project Summary (abstract)**

Will write last

**Intellectual Merit**

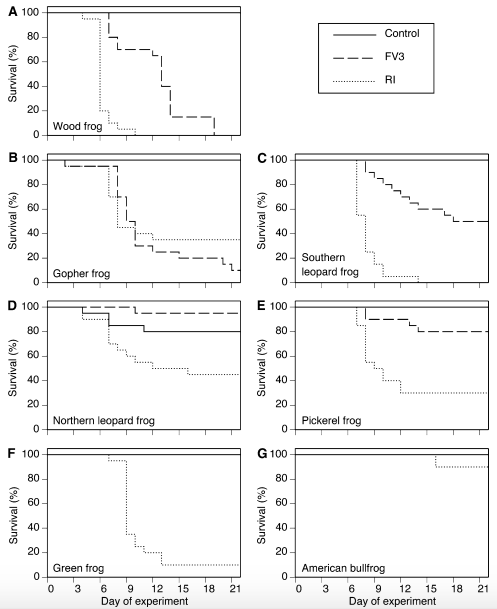
* Adds to the body of emerging infectious diseases, data for a region it has not been previously found in, adds to understanding of this particular disease (perhaps could apply to other systems?)
* The disease is killing frogs: could inform conservation decisions (protect vulnerable populations from disease-related die-offs), especially in Vermont
* Gaps in knowledge: how virus is being influenced; new locations of disease; warmer temperatures influencing disease prevalence?
* Our ability to predict the ramifications of these emerging infectious diseases depends on furthering our understanding of how diseases function in locations not previously found and changing environments, which are increasingly warming and developing (Schock et al. 2009).
* Ranavirus surveillance and population monitoring of natural communities are severely lacking, with most research concentrating on laboratory experiments (Gray and Chinchar 2015). For example, transmission routes have been tested in the lab, so it is known that ranavirus can be transmitted through direct contact, indirectly (i.e. water or sediment), and ingestion (i.e., cannibalism, predation).

**Introduction**

Viruses are considered to be the most ubiquitous and abundant organisms in the world, with an estimated 1031 viruses across the globe (Breitbart and Rohwer 2005). In fact, mathematical models have predicted that the diversity of viruses found in 1 kilogram of marine surface sediment is larger than the diversity of all reptiles on the planet (Breitbart et al. 2004). Not only do they serve as a reservoir for the greatest genetic diversity on Earth, they are important agents of mortality and are central in global geochemical cycles (Suttle 2005). Considering viral abundance, diversity, and ecological importance, surprisingly little is known about their biogeographical distributions, community structure, and ecological dynamics (Breitbart and Rohwer 2005).

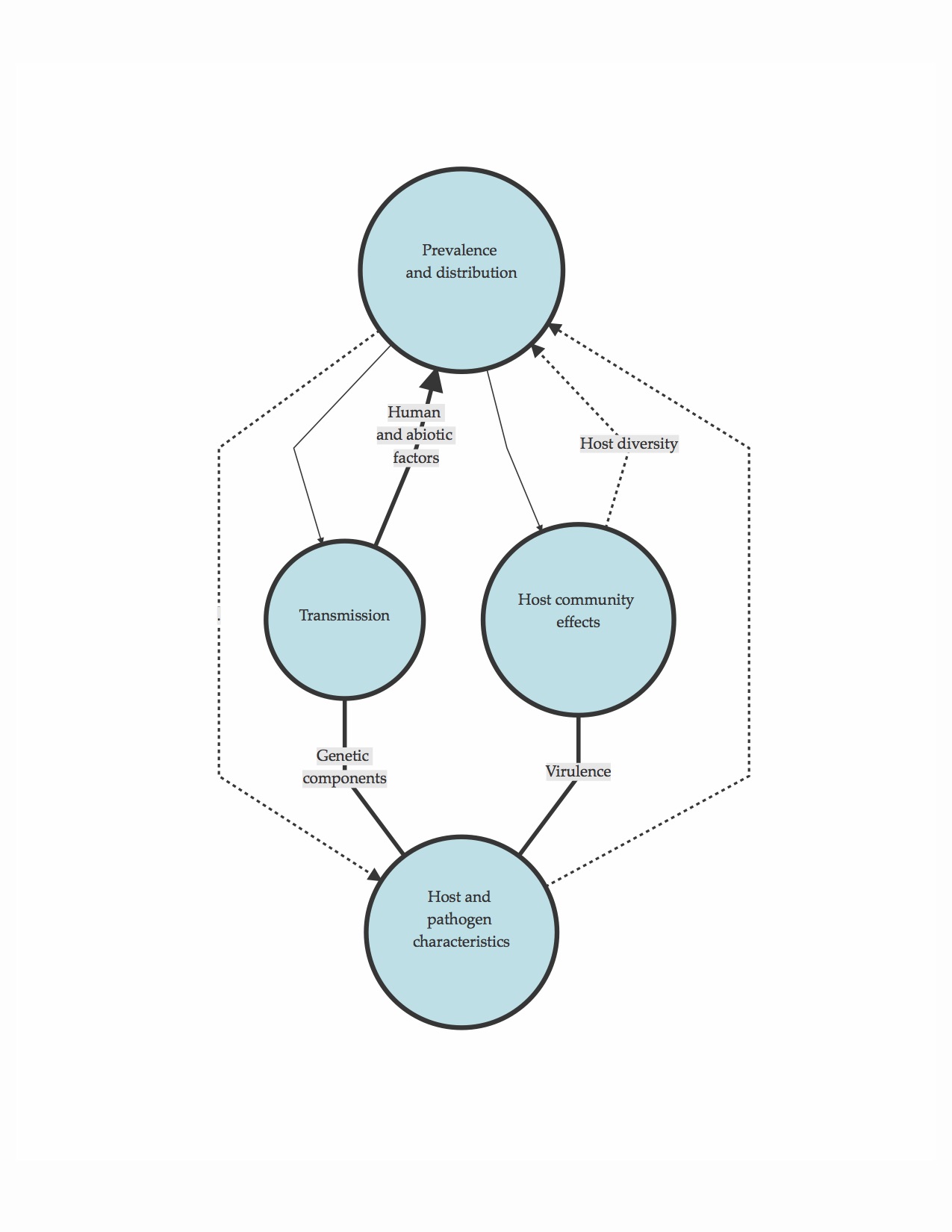
The genus *Ranavirus* encompasses a group of large, double-stranded DNA viruses (family Iridoviridae) and provides an example of the ubiquitous, yet relatively unknown, nature of viruses. These viruses infect a wide range of species, are transmitted through multiple routes, and have large, varying effects between species and populations, making them more general than most viruses; thus, they provide a unique system within viral studies. Ranaviruses have been documented in 32 countries on 6 continents and possess an extremely large host range, with the capability of infecting multiple species across classes (Chinchar and Waltzek 2014), specifically bony fish (Actinopterygii**)**, reptiles (Reptilia), and amphibians (Amphibia). Amphibians (Lissamphibia) are a diverse and abundant group of organisms that serve as indicators of environmental health and are a vital link in the food web (Wake 1991). However, amphibian populations worldwide have been in decline, and studies point to emerging infectious diseases as one of the major contributors (Gray and Chinchar 2015; Harp and Petranka 2006). Specifically, they have been shown to be particularly susceptible to ranaviral disease: one study reported that the most common cause of amphibian mortality events was infection by ranaviruses (Green et al. 2002). Other studies have reported that ranaviruses are resulting in population declines (Petranka et al. 2003; Teacher et al. 2010; Price et al. 2014; Wheelwright et al. 2014) and have the potential to cause local extinctions (Earl and Gray 2014).

Ranavirus infected wood frog larvae (image by Scott Smith)

There are 3 official *Ranavirus* species known to infect amphibians: *Frog virus 3* (FV3), *Ambystoma tigrinum virus* (ATV), and *Bohle iridovirus* (BIV). They have been reported in at least 105 amphibian species in 18 families (Duffus et al. 2015). Transmission of the virus can occur through several routes, including through water and substrate, direct contact, and ingestion of infected individuals (REF). Although the symptoms vary between host and pathogen species, host life stage and transmission route (REF), the viruses generally cause lethargy, internal and external hemorrhaging, swelling of the body and legs, and erratic swimming, with fatal cases involving necrosis in the liver, kidney, and spleen (REF). Additionally, the amphibian ranaviruses have a general trend in the timing of outbreaks, with most die-offs occurring rapidly in the mid to late-summer months (REF). Although these patterns can be observed, there are notable exceptions: the timing of outbreaks in certain species, like Bullfrogs, is much later in the summer (REF) and individuals can be asymptomatic (REF). Unsurprisingly, the mortality rate of infected individuals is also inconsistent (Figure 1).

**Figure 1** shows the difference in mortality between different anuran species and ranavirus strains. Tadpoles were exposed to virus-free media (Control), frog virus 3 (FV3), or an FV3-like isolate from a Georgia ranaculture facility (RI). For each treatment, n = 20 tadpoles (Hoverman et al. 2011)

The outcomes of these outbreaks can vary between species, populations and location, ranging from no apparent mortality to mass die-offs. Factors, such as host life stage, temperature, and anthropogenic influences (i.e. human visitation and pesticide use), have been implicated in this variation (REF). For example, the seasonal timing of ranavirus outbreaks coincides with both high temperatures and often the metamorphosis of amphibian larvae. However, studies that tested these influences found conflicting results. Arial et al. (2009) found that multiple amphibian ranavirus species replicated faster with increasing temperature up to a certain optimum, usually between 24°C and 28°C. Contrastingly, another study showed that salamander larvae reared at 10°C or 18°C experienced higher mortality after exposure to ATV compared to larvae reared at 26°C (Rojas et al. 2005), and proposed that the immune system could be suppressed in colder temperatures. Additionally, the process of larval metamorphosis involves natural immunosuppression (Rollins-Smith 1998; Carey et al. 1999); thus, it has been hypothesized that some of the variation in ranavirus outbreak timing could be explained by host life stage. However, although one study found a 1.7-fold increase in mortality of wood frog tadpoles exposed to ranavirus with increasing Gosner (1960) development stages (Warne et al. 2011), another found that metamorphosis was not always the most susceptible stage (Haislip et al. 2011). Finally, further environmental and anthropogenic factors may contribute to the variance in disease prevalence and host mortality, as well. Studies have shown increased probability of outbreaks in areas with cattle access (Gray et al. 2007; Greer and Collins 2008; Hoverman et al. 2012), low elevation (Gray et al. 2009b; Sutton et al. 2014), high elevation (Gahl and Calhoun 2010), and pesticides (Forson and Storfer 2006b; Kerby and Storfer 2009; Kerby et al. 2011).

Although there is strong evidence that ranavirus replication and the outcome of infection depend on the host and virus species, as well as other confounding environmental and anthropogenic factors (Speare and Smith 1992; Grant et al. 2003; Rojas et al. 2005; Ariel et al. 2009b), it is clear that ranaviruses have the potential to impact ectothermic vertebrate populations and can often trigger significant morbidity and mortality. Investigating the drivers of outbreak variation would not only further our understanding of generalist viruses and infectious disease ecology in general, but could inform amphibian conservation efforts, as well. Preliminary results from my current research indicate that ranaviruses are indeed present in the state of Vermont, an area where it has not yet been documented. My proposed dissertation research will focus on the ecological and anthropogenic variables associated with disease presence, the effects of these viruses on amphibian communities, how transmission is occurring between ponds and individuals, and the characteristics of both host and pathogen that influence host mortality (Figure 2).

**Figure 2** A path diagram detailing how my four questions interact

**Question I: What is the distribution, prevalence, and community structure of ranaviruses in amphibian communities of Vermont?**

**Objective I: a) I aim to evaluate the roles that temporal, spatial, environmental, and anthropogenic factors play in *Ranavirus* distribution and prevalence in natural amphibian populations of Vermont, and b) I aim to determine the diversity and abundance of *Ranavirus* species in the state of Vermont.**

**Hypotheses and Predictions**:

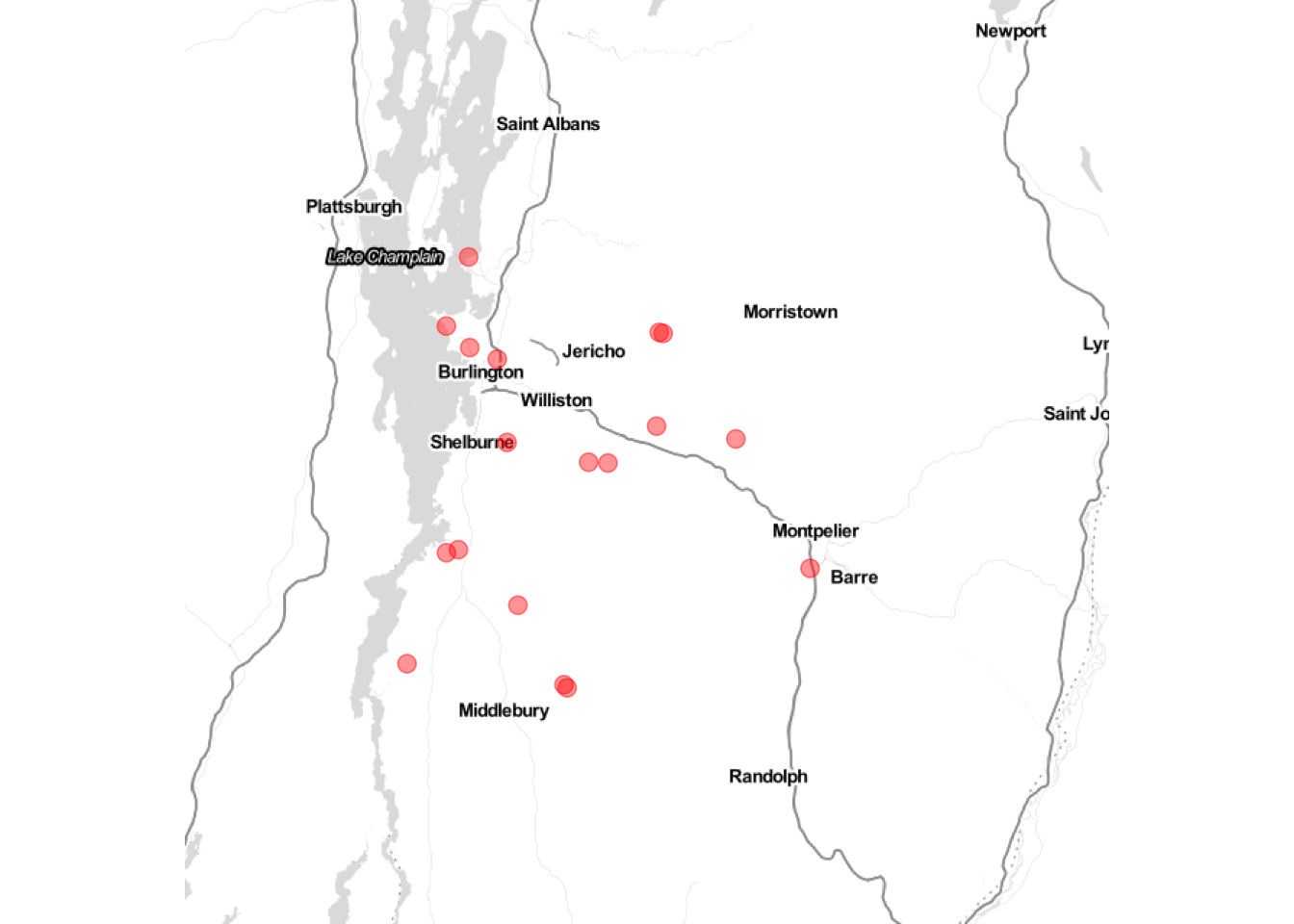
* I hypothesize that **ranaviruses are present in the state and that prevalence increases in areas with environmental and anthropogenic stressors.** I predictthere will be increased prevalence in sites closer to roads and agricultural areas and sites that have higher human visitation and fishing activity. I also predict that ponds with variable water throughout the season will have increased prevalence.
* I hypothesize that **viral diversity mimics the viral diversity of surrounding New England states**, and predict Frog-Virus 3 will be the primary species, with potentially multiple sub-strains.

**Background**

Surrounding Vermont, ranaviruses and associated mortality events have been found in amphibian communities of Maine (Gahl and Calhoun 2010), Massachusetts and New Hampshire (Green et al. 2002), New York (Brunner et al. 2011), New Jersey (Monson-Collar et al. 2013), and Quebec (Paetow et al. 2011) and Ontario (Greer et al. 2005), Canada. Prevalence has not been reported for most studies. Quebec found ranavirus prevalence at 85 percent (Paetow et al. 2011), but another study found in 2012 that 100 percent of 18 ponds across 10 states on the United States eastern coast had ranavirus-infected frogs (Crespi et al. 2015). Although a 1968 New York study found ranavirus in the liver of Northern Leopard Frogs (*Lithobates pipiens*) purchased from a business in Vermont (Clark et al. 1968), no ranavirus surveys have been conducted in the natural amphibian communities of the state. Therefore, ranavirus distribution, prevalence, and its associated factors are virtually unknown. More studies are needed to understand why ranaviruses emerge in particular areas (Gray and Chinchar 2015), so the fact that it has not been detected here, as of yet, makes research in the state warranted. Thus, I will use a predictive modeling approach to determine the distribution of ranaviruses and the variables associated with ranavirus occurrence.

The diversity of ranavirus species and strains in New England and Canada is also not well known, as strains are not reported in most of the literature. However, after genetic analyses, a study in New York found FV3-like ranaviruses present and studies in Ontario and Quebec observed ranavirus sequences were a 98 and 100 percent match to FV3, respectively. Frog Virus 3 (FV3) appears to be the most abundant ranavirus species in New England, but there are also dozens of strains that have less than 98% genetic similarity (He et al. 2002; Jancovich et al. 2015; Lei et al. 2012; Morrison et al. 2014). To estimate ranavirus diversity in Vermont, I will use sequencing and phylogenetic methods on the amphibian tissue samples I collect.

**Approach**

*Sampling Design*

In the summer of 2016, I aimed to collect amphibian tissue samples across sites in northwestern Vermont (Figure 3). Sampling would occur at each site, once every other week, from mid-May to August in order to increase the probability of witnessing a mass mortality event and because the sensitivity of PCR using non-lethal tissue samples peaks around 12 days post-virus exposure (Greer and Collins 2007). The potential sites were chosen with James Andrews, the state herpetologist of Vermont. Sites with predicted high amphibian abundance, estimated from Mr. Andrews’ prior survey experience, ideal amphibian conditions (i.e. shallow water and emergent vegetation), assessed either from prior knowledge or from Google Maps satellite view, and those that could reasonably be visited once every other week were selected as contenders. The consistently high amphibian abundance throughout the summer at these sites was considered because an increase in the amount of samples collected would increase the probability of disease detection. A total of 18 sites met the requirements, were chosen as final sampling locations, and were visited a total of seven sampling periods throughout the summer.

**Figure 3** A map of the 18 sites in northwestern Vermont where amphibian tissue was collected every other week from mid May to August 2016.

*Sampling Techniques*

Previous estimates indicate that for 95% confidence in disease detection and at 5% disease prevalence, tissue should be collected from a sample size of at least 20 individuals (Gray et al. 2015). Therefore, samples were taken from a maximum of 30 individuals at a particular site or after an allotted amount of time, to standardize sampling effort. To obtain a total amphibian search time of an hour and a half ‘person hours’, either two people searched for 45 minutes or three people searched for 30 minutes. All non-endangered species of amphibians at various life stages were collected and placed in either individual plastic bags of distilled water, if a salamander or larval frog, or individual plastic containers, if an adult frog, to limit cross-contamination. Additionally, different pairs of powder-free nitrile gloves were worn while handling each individual. Although lethal methods of sample collection (i.e. liver samples) detect disease more precisely, being able to accurately determine the distribution and prevalence of the disease in Vermont required too large of sample sizes to warrant lethal sampling. Therefore, tail tissue was collected from salamanders and larval frogs, and toe tissue was obtained from adult frogs. Tail collection involved pressing the flat side of a ruler onto the tip of the tail, triggering the natural predator defense of tail autonomy, which minimizes blood loss. Using surgical scissors, one toe was collected from an adult frog per sampling week. If the individual was recaptured, non-adjacent toes were clipped. Forceps and scissors were disinfected between uses with 10% bleach, and waders, boots, and nets were disinfected between sites using 3% bleach, according to the Northeast Partners in Amphibian and Reptile Conservation disinfection protocol (NEPARC 2014). Tissue was stored in 1.5 mL tubes of 90% ethanol in a -20° freezer. A total of 1,822 tissue samples from a total of 10 amphibian species were collected throughout the summer.

*Testing for virus*

DNA was extracted using the Omega Bio-Tek E.Z.N.A. Tissue DNA kit and protocol. I tested for presence of the virus using real-time quantitative PCR (qPCR) because studies have shown that qPCR methods are more sensitive to virus detection when compared to PCR, especially when pathogen concentration is low (Monson-Collar et al. 2013). Real-time quantitative PCR was performed in MicroAmp optical 96-well reaction plates using StepOnePlus Real-Time PCR Systems (Applied Biosystems). The 10uL reactions consisted of 2 uL of 10ng DNA, 5 uL 1X Sso Advanced SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA, USA), 0.25 uL of forward and reverse primers, and 2.75 uL of nuclease-free water. Primers RVMCPKim3\_F (5’-TAA- CACGGCATACCTGGACG-3’) and (5’-GATGAGATCGCTG- GTGTTGC-3’) RVMCPKim3\_R (Kimble et al*.* 2014) were used to amplify a 97 bp region of the major capsid protein. The thermal profile consisted of 2 minutes at 95°C, then 35 cycles of 95°C for 20 seconds followed by 65.1°C for 20 seconds. Each individual sample was run in duplicate with four serially diluted standards from 10 to 106 viral copies constructed from cultured Frog Virus-3 (FV3), and a negative control using sterile H2O. qPCR standards were created from cultured FV3 from an outbreak in wild adult Northern Leopard Frogs (*Lithobates clamitans*) from Illinois and obtained from Dr. James Julian at the University of Pennsylvania Altoona. Based on the standard curve, a reaction was considered positive when Ct < 34.0.

*Summer 2017*

If ranavirus is present in Vermont and does not show significant variation in prevalence between weeks, I plan on expanding the number of sites I collect samples from, but reducing my sampling periods to three or four times during the summer. The decrease in sampling periods is due to the fact that mass mortality events were not witnessed during the previous summer. Therefore, reduced sampling may still capture representative prevalence throughout the summer. I plan to record amphibian mass data and environmental data, such as water and soil quality, from sites each time I collect samples.

*Environmental DNA*

Additionally, I will obtain environmental DNA (eDNA) to test for the presence of ranavirus, in order to approximate how many ranavirus virions are present in the water and sediment. eDNA methods are not common in this system, yet one study showed that there is a strong relationship between eDNA viral titres in the water and titres in larval tissue (Hall et al. 2016). Also, to the best of my knowledge, no study has extracted ranavirus from sediments, although a study showed the virus can persist in soil for 30 to 48 days at 4°C (Nazir et al. 2012). I will collect water eDNA using a filter, hand pump, and vacuum flask, and soil eDNA will be collected using a hand corer. The water filter and soil DNA will be extracted using kits. I will test whether this is the case for my tissue samples and both water and sediment eDNA samples, and can determine whether water or sediment contains greater viral titres.

*Data Analysis: Prevalence*

To address my first hypothesis, virus prevalence can be estimated by dividing the number of cases by the population size. In my case, I will estimate prevalence at a given site and time by dividing the number of infected individuals by the total sample size. I will also be able to compare prevalence estimates between species. To determine whether there are significant differences in prevalence between species and sites, I can use Pearson’s chi-squared test. To estimate factors associated with ranavirus occurrence, I can use logistic regression.

*Data Analysis: Predicted distribution and potential predictor variables*

In order to attempt to understand what is driving the presence of the virus in certain locations, I can first determine if the means of bioclimatic variables and elevation are different between ponds where ranavirus occurred at some point throughout the summer and those where it did not, using simple one-way analyses of variance (ANOVA). Next, I can use either these coarse BioClim data layers or finer resolution Vermont-specific data layers to create a predicted distribution of the presence of ranavirus in the state. I can start with all 21 predictor layers, determine which are strongly influencing the model, and reduce the number of variables. Since the number of predictor variables may still be more than the number of sites, I can compare the full model to a model that uses a raster PCA of the data layers. Using AIC model evaluation, I can determine the best predictive model and identify potential disease ‘hotspots,’ or areas with a high probability of presence. Next, if a relatively decent model can be constructed for the small area of Vermont, I can conduct a meta-analysis using latitude and longitude points from other studies to form a model for a larger area. Careful selection of other coordinates will need to be made, in order to avoid variance between years and potentially months. Using the driving predictor variables, I can also create a regression tree to see exactly the values that are driving separation between present and absent sites. Additionally, because I plan on collecting environmental information at each site and time point this summer, I will be able to compare water and soil quality values to prevalence numbers. With these types of data, a linear regression will be the best method to determine if there is a relationship between those specific environmental conditions and disease prevalence.

*Data Analysis: Virus diversity and invasion history*

The amplified ranavirus-positive DNA will also be Sanger sequenced at the Advanced Genome Technologies Core Facilities at the University of Vermont. With the sequenced amplified region of the major capsid protein, I can address my second hypothesis and determine which (and potentially how many) species of the six official viruses are present in the Vermont amphibians I collected. In the future, I would like to use next generation sequencing to sequence the whole genome of positive samples and determine which sub-strains are present, as well. Using that collection of ranavirus sub-strains, and potentially sequences along state borders, I can construct a virus phylogenetic tree. From there, I can use a neighbor-joining tree analysis and haplotype distribution to explore the invasion history of the virus into the state. These particular methods have not been used before in the ranavirus system, yet have the potential to allow us to better understand how, or if, the virus is spanning large areas and the best method of limiting its spread.

**Expected Results and Implications**

I expect to obtain two main results from this first question. Firstly, ranavirus has not been documented in natural communities of Vermont, and this is the first large survey of the state. I will be able to present the first instance of the virus, its predicted distribution and potential disease ‘hotspots’ in Vermont, as well as its prevalence and diversity in the sites I sampled. Furthermore, to the best of my knowledge, the predicted distribution of the virus has not been published. For my second main result, I can use a combination of spatial and phylogenetic analyses to present an invasion history of the disease, first locally in the state and then broader-scaled. This method has not been used in this disease system, but has the potential indicate if expansion is human-mediated.

**Question II: What are the effects of ranaviruses in amphibian communities of Vermont?**

**Objective II: I aim to determine a) whether there is a relationship between species diversity and disease presence and virulence and b) if and why disease severity varies across sites in Vermont.**

**Hypotheses and Predictions**:

* I hypothesize that **ranavirus outbreaks are causing mass mortalities and reducing amphibian species diversity, genetic diversity, and species abundance**. I predict that sites with high ranavirus prevalence will have reduced amphibian species richness, evenness, abundance and genetic diversity.
* I hypothesize that **host community and genetic structure cause variation in disease severity** between amphibian populations and communities and predict that communities and populations with low genetic diversity, richness, and evenness will have increased ranavirus virulence.

**Background**

Biodiversity has repeatedly been shown empirically and theoretically to be associated with community functioning and stability (Shear McCann 2000; OTHER REFS). Higher biodiversity, and hence many weakly interacting species, leads to increased ecological stability by dampening strong, and potentially destabilizing, consumer-resource interactions (Shear McCann 2000). In fact, the removal of just one species can lead to drastic community changes (Shear McCann 2000). Ranaviruses have the capability of causing mass mortalities and have the potential to cause local extinctions (Earl and Gray 2014), and therefore, determining the effects of these viruses on the community structure and diversity is important. In addition, more stable, biodiverse communities may be able to handle stressors, such as disease, better than unstable communities (REF). If so, estimating community diversity may enhance our understanding of how communities will be able to respond to ranaviruses.

In addition, there could be a relationship between disease and genetic diversity. Estimating a population’s genetic diversity before and after high disease prevalence may shed light on whether mass mortalities are occurring. If selection, in the form of disease, is causing a population bottleneck, we may be able to pick up differences in population-level allelic diversity and heterozygosity. Similarly to biodiversity, genetic diversity may also influence a community’s response to disease, and with increased genetic diversity, a population can better respond and adapt to a selective pressure.

The variation in virulence, or host fitness consequences, across species and populations is another factor that could influence how communities respond to these viruses. Virulence can be measured as host mortality, and disease severity and mortality have been shown to strongly correlate with viral load (REF). Thus I will use viral load quantification as a measure of disease severity. Not only do I aim to determine whether an association exists between virus prevalence and host diversity, but also if disease severity can either affect or be explained by differences in host genetic and species diversity.

**Approach**

*Viral load quantification*

Viral load will be used as an estimate of disease severity. It will first be estimated for all samples by comparing the cycle threshold (the crossing point of the amplification curve with the preset threshold of fluorescence detection) of the sample to the standard curve. Using Nanodrop, the amount of DNA in each sample was estimated and then diluted to a concentration within an order of magnitude to each other (less than 100 ng/uL). Using this method, a rough estimate of viral copy number for each positive sample is calculated, and the resulting estimates can then be compared. In the future, to obtain more precise calculations of viral copy number, the positive samples will be run again in triplicate with an amphibian reference gene. This allows for the amount of host DNA in the reaction to normalize the qPCR data. Amphibian genes such as ribosomal protein L7 (RPL-7), and 18S ribosomal RNA have been used as reference genes in the past (Zhang and Hu 2006), but I will have to perform tests to determine which gene would work best with my qPCR assay.

*Calculating species and genetic diversity*

To measure species diversity, a diversity index that summarizes the number of species and their relative abundances in a community will be calculated. Simpson's diversity index (*D*) is often used as an estimate of species diversity (Equation 1; Simpson 1949).

(*Equation 1*)

In addition, the uniformity of species abundances, or evenness, can be estimated with the Probability of Interspecific Encounter (PIE; Equation 2; Hurlbert 1971).

(*Equation 2*)

To calculate genetic diversity within a host population, ideally I would like to use next generation sequencing to obtain whole genome diversity estimates. However, another less expensive option to obtain genetic diversity would be to use microsatellite markers. A previous study has estimated genetic diversity for the population of one amphibian host that was infected with ranavirus and found that genetic diversity was lower in ranavirus present sites (REF). I would like to estimate genetic diversity and structure for multiple host populations, or each species that was infected with ranavirus at some point throughout the summer. Common ways of estimating genetic diversity are using heterozygosity and allelic richness and diversity at certain loci or SNPs (Allendorf 1986). Heterozygosity can be used as a measure of the capability of a population to respond to selection immediately after a bottleneck (i.e. disease), while allelic diversity determines the capacity of a population to respond to long-term selection over numerous generations (REF). In addition, I would like to compare genetic variability within and between populations using FST.

*Analyses*

To determine whether there is a difference in species diversity between sites that have ranavirus infected species and those that do not I will perform t-tests. I can also use hierarchical clustering analyses and non-metric multidimensional scaling (NMDS) to further explore the data. To determine whether there is a relationship between virus prevalence/severity and diversity, I will use generalized linear models to perform regressions. With this method, I can determine how much of the variance in the species and genetic diversity can be explained by disease prevalence and viral load, and vice versa. Additionally, I would like to perform a regression on disease prevalence/viral load and abiotic measurements to see if disease severity can be explained by these abiotic factors (i.e. water pH). To determine whether there is a relationship between genetic differentiation and disease severity, I can apply the same method using FST.

**Expected Results and Implications**

Determining whether community assemblages are having an effect on disease occurrence (or vice-versa) is a critical component of disease ecology that has often been overlooked in the ranavirus system. Few ranavirus studies have collected samples from amphibian communities, often focusing on the effects of disease on the susceptible species *Lithobates sylvaticus* (Wood Frogs). By determining the effects of ranavirus on the amphibian communities of Vermont, I aim to determine whether ranavirus-associated mass mortalities are occurring and if there is a relationship between disease severity and host species diversity.

If mass mortality events were occurring at diseased sites, I would expect to find differences in species diversity between disease present and disease absent sites or a relationship between disease severity and host diversity. This would be due to susceptible species dying either to local extinction or to smaller numbers, which would impact relative abundance. A significant decrease in total abundance after high ranavirus prevalence could indicate mortality is occurring, as well.

However, if I do not observe these differences or relationships, it could be due to two reasons. First, perhaps mortality events occurred after I stopped sampling in August. I intend to lengthen my sampling season this summer to address this. Next, perhaps ranavirus in these Vermont communities are not causing mortality. There have been reports of ranavirus-positive amphibians without notable disease or mortality (Duffus et al. 2008; Duffus et al. 2015; Greer et al. 2009). However, sublethal infections have been shown to impact fitness-related traits, such as growth and development (Echaubard et al. 2010). If Vermont amphibians are showing tolerance to the disease, perhaps testing fitness could determine whether the virus is affecting them.

Another unexpected alternative result would be if higher genetic and/or biodiversity is shown to be associated with higher disease severity. A possible explanation that would need investigating would be that host species diversity might breed pathogen genetic diversity. This could increase the probability of ‘accidental’ or purposeful high virulence. Another explanation could be that populations with higher genetic diversity could be experiencing higher rates of migration, which could be exposing them to ranavirus strains more frequently.

**Question III: How do amphibian community dynamics influence ranavirus transmission and survival?**

**Objective III: I aim to determine a) how different species interactions influence virus transmission and b) how host density influences the ability of viruses to persist outside of their host.**

**Hypotheses and Predictions**:

* I hypothesize that in addition to environmental virus exposure from water and soil, the **frequency and routes of interaction between species are influencing virus transmission** in amphibian communities of Vermont. I predict ponds where species are coming into contact with each other more, such as predation, necrophagy, and mating interactions, will have increased virus transmission.
* I hypothesize that **pathogens are constrained by a survival-reproduction tradeoff** and maximize their fitness by employing different life history strategies under different environmental circumstances. I predict that increasing the length of time a pathogen must spend in the environment until it encounters a host will decrease pathogen reproductive fitness because it will invest more energy into surviving outside of the host.

**Background**

Ranavirus transmission can occur between individuals through direct contact

- How is it transmitted between individuals? (more detail; ingestion has bigger dose! Thus species network is warranted)

- How do humans influence transmission (other ways to go between sites)?

- Previous eDNA research

- Host-switching behavior (general/ranavirus) ## persistence outside amphibians

- Experimental evolution: viruses dominate our planet and their evolution is a broad and applied field that can be studied in the real time

-life history trade off: survival-reproduction

- SIR models (equations); previous ranavirus SIR models (Duffus 2009)

- previous studies of Wasik and De Paepe

- conflicts with curse of the pharaoh (Bonhoeffer et al. 1996)

- plaques

- phages/coliphages: variety in tails, lytic vs lysogenic, T-even vs T odd

“Life-history theory predicts that traits for survival and reproduction cannot be simultaneously maximized in evolving populations; selection for improved between-host survival during transmission may lead to evolution of decreased within-host reproduction”

“parasites with extreme survival will evolve mechanisms for this trait that should often compromise growth ability.” (Wasik et al. 2015) -> pleiotropy!

*PhiX174 belongs to the Microviridae family of bacteriophages (*[***12***](http://aem.asm.org/content/76/21/7310.full#ref-12)*). It is a small, icosahedral, nontailed virus with a circular single-stranded DNA*

*the replication of PhiX174 DNA can be supported by different Escherichia coli strains and distantly related bacteria, such as Pseudomonas aeruginosa (*[***28***](http://aem.asm.org/content/76/21/7310.full#ref-28)*). These findings imply that the limiting step for PhiX174 infection is entry and not replication or lysis. The commonly used PhiX174 host is the laboratory-derived strain E. coli C, which has a specific rough lipopolysaccharide (LPS) recognized as the receptor (*[***13***](http://aem.asm.org/content/76/21/7310.full#ref-13)*).*

SIR models are not used frequently in the ranavirus system, and no published studies exist on ranavirus transmission rates or dynamics in wild populations (REF). Therefore, to address my first question, I will model species interactions that I believe are important in transmission dynamics and variable between ponds. From there, I will apply the different predicted contact rates to alternate SIR models, use species and environmental viral load data, and determine how closely the SIR predictions of timing and prevalence match with field observations. More generally, I would like to investigate how host density influences viruses’ ability to persist outside their host. To address this, I will also conduct an experiment using viruses that infect bacteria (bacteriophages, or phages), to a) determine if a survival-reproduction trade-off still remains in multiple phage strains after altering timing-of-transmission and b) to infer the phenotype-genotype associations.

**Approach: SIR modeling**

To begin, I will model the number of susceptible, infected, and recovered individuals for one species and incorporate its interactions with other species and the environmental virions into the model.

*Parameters*

Using my species richness and abundance data from summer sampling, as well as a matrix of potential interactions (including predation, mating, necrophagy), I will use packages in R (R Core Team 2017), such as enaR (REF), foodweb (REF), and/or bipartite (REF), to perform ecological network analyses. I can use model output, such as edge weights, to inform the transmission and contact rate parameters of alternate SIR models. In addition, I can use average viral loads of each species and environmental DNA measurements of viral load in the water and substrate at each site, as well as species-specific mortality, infection, and recovery rates (using my data and data from the literature) as parameters of the model.

*Model construction*

The creation and implementation of the SIR models will also be conducted in R. I will start with the EpiModel package (REF); and with increasing complexity, I will create ordinary differential equations and SIR models outside of the package, which is the more flexible option. After obtaining predictions of prevalence and timing from the SIR models, I can compare them to field observations of prevalence and timing. In addition, by changing parameters, I can observe which factors are strongly influencing ranavirus transmission and the amount of infected individuals.

**Approach: Experimental evolution experiment**

I would like to test how increasing timing-of-transmission influences the reproduction-survival trade off and if the effects differ between coliphages. The following experimental design was informed by the previous studies of De Paepe and Taddei (2006) and Wasik et al. (2014).

*Phages*

I selected 5 well-characterized lytic DNA coliphages that were shown by De Paepe and Taddei (2006) to have differential survival and reproduction: **φ**X174 (ssDNA; family Microviridae), T5 (dsDNA; family Siphoviridae), T4 (dsDNA; family Myoviridae), PRD1 (dsDNA; family Tectiviridae), and T7 (dsDNA; family Podoviridae) will be used in this study. As they are all coliphages, I can keep the host constant by using *E. coli* for all phages. In addition, phage stability is very dependent on salt concentration and osmotic pressure, so I will use the standard high-salt LB broth instead of medium that could potentially favor one phage over another (De Paepe and Taddei 2006).

*Experimental Design*

A single ancestral genotype of each phage will originate 8 lineages evolved in either ordinary cell culture passage (n = 4 lineages) or in delayed cell culture passage (n = 4 lineages), totaling 40 experimental units. To determine how long ‘ordinary’ and ‘delayed’ passages will be for each phage, I will conduct a pilot study to observe at what time populations reach a stationary density (ordinary), and then double that time (delayed). The lineages will be observed for 30 passages and frozen every passage to obtain an ‘evolutionary record.’ Fecundity and extracellular survival will be measured for each lineage every 10 passages.

*Sequencing*

At the end of the experiment, each lineage will have its whole genome sequenced every 3 passages (using PCR and Sanger sequencing of viral RNA), and the number of non-synonymous and synonymous substitutions will be compared between the ordinary or delayed passage treatments.

*Fecundity and survival measurements*

Fecundity will be used as a proxy for reproductive fitness and will be measured as delta-log10 titer. First, I will conduct 6-h growth productivity assays of phage populations. Approximately 104 viruses of each experimental lineage (3 replicates) will be allowed to infect replicate 25 cm2 flasks. The phages will be counted using plaque assays, and these data will be log10 transformed and divided by the population titer (104 plaque forming units) to obtain delta-log10 titer. Extracellular survival will be measured by placing 106 pfu/mL of virus in 10 mL of cell-free DMEM and calculating the remaining titer with plaque assays at 5 time points (6, 12, 24, 36, and 48 h post incubation).

*Statistical analyses*

To determine whether there is a strong relationship between survival and reproduction, I will perform a linear regression model. I can conduct an ANCOVA to determine whether the slopes (i.e. tradeoffs) are different between delayed and ordinary treatments and between the different phages. If slopes are different between treatments, I can perform a piecewise regression to estimate the ‘breakpoint’ (or break in the tradeoff). This can be done with the segmented package in R (Muggeo 2008).

**Expected Results and Implications**

Based on how similar the predictions and observations are, we can determine how much influence species interaction has on transmission and hence prevalence of ranavirus.

-First transmission network/SIR model in system

-No published studies on ranavirus transmission rates or dynamics in wild populations

-Will either see a trade-off or a break in trade-off at a certain point

-If all slopes are the same across phage species, indicates there is indeed a trade off; if slopes are different, perhaps it is more complex than that

-Expected results experimental evolution: more non-synonymous substitutions in delayed transmission lineages because signature of selection

one substitution could be enough to increase performance (pleitropy)

**Question IV: Are there host characteristics that influence their susceptibility, resilience, or tolerance to ranaviruses?**

**Objective III: I aim to determine a) how different species interactions influence virus transmission and b) how host density influences the ability of viruses to persist outside of their host.**

**Hypotheses and Predictions**:

* I hypothesize that **amphibian populations that have been previously exposed to the disease are locally adapted** and will have higher ranavirus tolerance
* I hypothesize that **asymptomatic and surviving tadpoles have a higher tolerance of the virus.** I predict they will have adaptive alleles and show higher differential gene expression in immune-related genes.

**Background**

- Define host susceptibility, tolerance, and resistance

- Has differential gene expression been done in this system?

- What is known about ranavirus genomics? 11 currently published and completely sequenced (p62)

- Genetic basis of pathogen

- Differences between populations that have been exposed to it before

- tolerance vs resistance

- why choose Wood Frog? It is found across Vermont; has been shown to be infected in Vermont; has differential mortality (Hoverman et al. 2012); many eggs early in the season

I will use common garden experiment to determine the genetic component of ranavirus symptom variation

-I will be attending a training course

I can approach this question with different methods: population genetics, experimentation, and genomics (next gen sequencing of field samples and experiment samples).

-set other factors constant

-mock data set- glm model, binomial link; random =eggs; response 0/1

I will also use population genomics and single nucleotide polymorphism (SNP) data to determine if ranavirus is acting selectively on particular loci in *Lithobates clamitans.* Population genomics uses SNPs to analyze genetic differentiation between groups, determine population structure, elucidate demographic history, and test for signatures of selection (REF). SNPs are genetic markers whose precise identification is a prerequisite for genome-wide association studies, which aim to find an association between a trait (i.e. health status) and markers (REF). Differential gene expression analysis uses …

To determine whether the genetic composition of the host influences the response to ranaviruses, I would like to experimentally infect tadpoles and measure allelic differences, as well as differences in gene expression.

**Approach**

5 ponds and 5 ponds: 4 egg masses, 10 tadpoles = 400 tadpoles

keep track of egg mass; treat as block; if possible, one infected and one uninfected from each egg mass

10 tadpoles per egg mass to begin with;

glm multinomial link?

*Collection and rearing*

I will collect 5 egg masses from 8 Wood Frog (*Lithobates sylvaticus*) populations, with 3 populations from sites where ranavirus has been recorded in the past two years and 3 where the virus has not been recorded. The egg masses and larvae will be reared in laboratory conditions at 23°C and a 12:12 hour day:night photoperiod (Relyea & Werner 1999; Hoverman et al. 2010; Hoverman et al. 2011). To minimize the probability of egg mass ranavirus infection, I will collect the eggs in early spring and test a subsample of eggs and tadpoles for virus with quantitative PCR. In addition, I will test the water of the site at the time of egg collection for presence of the virus. Eight tadpoles from each egg mass will be randomly assigned to either a control group or experimental infection treatment, totaling 240 experimental units. The tadpoles will be housed in individual plastic containers and the de-chlorinated aged tap water will be changed every 3 days (Hoverman et al. 2010). Because ranavirus susceptibility changes over life stages (Tweedell and Granoff 1968), larval development will be standardized, and tadpoles at Gosner (1970) stage 30 will be used in the experiment.

*Experimental infection*

Frog Virus 3 ranavirus isolates can be cultured at the University of Georgia Veterinary Diagnostic and Investigational Laboratory (VDIL), sent to the University of Vermont, and stored at -70°C until use. Doses between 102 and 106 plaque forming units (PFUs) of the virus have been shown to be sufficient to induce sublethal effects or morbidity in tadpoles (Tweedell and Granoff 1968, Pearman et al. 2004, Pearman and Garner 2005, Morales and Robert 2007). I will orally inoculate the tadpoles with 103 PFUs of the virus isolate suspended in 10 ul of Eagle’s minimum essential media (MEM Eagle Sigma-Aldrich, Seelze, Germany) using a pipette. The control group will be exposed to the same volume of virus-free media (MEM).

*Tissue collection*

Before experimental infection begins, 5 individuals from each egg mass will be sacrificed and necropsied. Liver and kidney tissue will be collected and stored in -80°C until DNA and RNA extraction. RNA will be used for RNA-seq analysis (described below) and DNA will be used to quantify the infection using quantitative PCR methods (previously described). From previous studies, green frog survival seems to drastically decrease around day 8 post-infection and then incrementally decrease until day 13 (Figure). On the first day that symptoms are observed, 5 symptomatic individuals, 5 asymptomatic individuals, and 5 control individuals will be sacrificed and have their liver and kidney tissue removed. I will then collect tissue from 5 control, moribund, and asymptomatic individuals every three days until day 14. Numbers or time points will be adjusted if there are not enough asymptomatic individuals. All surviving tadpoles will be euthanized and their tissue will be collected, as well.

*RNA-Seq and differential expression analysis*

Poly(A) selection will be used to obtain mRNA from the samples, which will then used to prepare the cDNA libraries. The libraries will be sequenced at the University of Vermont’s Advanced Genome Technologies Core Facility on Illumina HiSeq 3000 to produce short paired-end reads. Each sample will be run simultaneously on different lanes to reduce lane effects. I will trim the data of adaptors, which indicate the sample, and low quality reads using Trimmomatic. The genome of *Lithobates sylvaticus* is already published (Robertson and Cornman 2014), so I will use that as a reference genome and map my reads to it. Next, I will extract the read count information, which is the number of reads that mapped to each contig. To test for differential gene expression between control, symptomatic, and asymptomatic individuals, I will use the R package ‘DESeq2’ (Love et al. 2014).

*Population genomics and SNP data*

To extract SNPs (variable basepair positions of the transcript) and predict the posterior probabilities of genotypes, I will use the reads2snps program (REF), executable in Terminal. The program vcftools (REF) will be used to further filter SNP sites and genotypes, using metrics such as quality thresholds, minor allele frequency, biallelic SNPs, and maximum missing data. I will calculate allele frequencies and Fst (the differentiation between populations) using the filtered SNPs and perform discriminant and principal component analyses in R.

If there is funding and time, I can compare the alleles of these loci to my field sampled individuals that survived in infected ponds until the end of collection to see if they have the same alleles.

*Fitness experiments*

I will compare length and weight of surviving infected individuals to the control group to determine if there are negative impacts on growth associated with sub-lethal infection.

**Expected Results and Implications**

Previous studies have shown that individuals activate an immune response in reaction to the infection (REF); however, these results will be the first in the ranavirus system to use experimental infection to determine the genetic influence of ranavirus symptom variation. I expect to see individuals from ponds where ranavirus infection has been previously observed to show increased survival. I expect this because ranavirus has been shown to have increased virulence for naïve populations (REF). However, if I observe the opposite, in that ‘naïve’ populations show higher survival, perhaps they show resistance, which is the reason why the infection was not found in the first place. Another possible explanation could be that populations that had experienced infection beforehand suffered a bottleneck and had less genetic diversity to begin with. These can be further explored with the genetic analyses.

**Conclusions:**

* Synthesize predicted results
* How do they connect?
* How do they answer big question?

**Broader Impacts**:

* Conservation
* Working with the State Parks
* Vermont Herp Atlas
* Other database that woman emailed me about: ANR?
* Development as graduate student

**References**:

Bergh, Ø. et al. (1989) High abundance of viruses found in aquatic environments. Nature 340, 467–468

Breitbart, M. et al. (2004) Diversity and population structure of a nearshore marine sediment viral community. Proc. R. Soc. Lond. B. Biol. Sci. 271, 565–574

Brunner JL, Barnet KE, Gosier CJ et al (2011) Ranavirus infection in die-offs of vernal pool amphibians in New York, USA. Herpetol Rev 42:76–79

Crespi EJ, LJ Rissler, NM Mattheus, K Engbrecht, SI Duncan, T Seaborn, EM Hall, JD Peterson, JL Brunner. 2015. Integr Comp Biol. **55**: 602-617.

Gahl MK, Calhoun AJK (2010) The role of multiple stressors in ranavirus-caused amphibian mortalities in Acadia national park wetlands. Can J Zool 88:108–121

Gosner KL (1960) A simplified table for staging anuran embryos and larvae with notes on identi- fication. Herpetologica 16:183–190

Gray et al. 2015: Ranavirus Book - Design and Analysis of Ranavirus Studies: Surveillance and Assessing Risk

Green DE, Converse KA, Schrader AK (2002) Epizootiology of sixty-four amphibian morbidity and mortality events in the USA, 1996-2001. Ann N Y Acad Sci 969:323–339

Greer A.L., J.P. Collins. 2007. Sensitivity of a diagnostic test for amphibian ranavirus varies with sampling protocol. Journal of Wildlife Diseases **43**:525-532

He JG, Lu L, Deng M, He HH, Weng SP, Wang XH, Zhou SY, Long QX, Wang XZ, Chan SM (2002) Sequence analysis of the complete genome of an iridovirus isolated from the tiger frog. Virology 292:185–197

Hurlbert, S.H., 1971, The nonconcept of species diversity: a critique and alternative

parameters. Ecology, v. 52, p. 577-586. (On explicit calculation of rarefaction

richness and derivation of Hurlbert’s PIE measure of evenness.)

Lei XY, Ou T, Zhu RL, Zhang QY (2012) Sequencing and analysis of the complete genome of Rana grylio virus (RGV). Arch Virol 157:1559–1564

Love MI, W. Huber and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, **15**:550. doi: [10.1186/s13059-014-0550-8](http://doi.org/10.1186/s13059-014-0550-8).

Monson-Collar K, Hazard L, Dolcemascolo P (2013) A Ranavirus-related mortality recent and the first report of Ranavirus in New Jersey. Herpetol Rev 44:263–265

Morrison EA, Garner S, Echaubard P, Lesbarreres D, Kyle CJ, Brunetti CR (2014) Complete

genome analysis of a frog virus 3 (FV3) isolate and sequence comparison with isolates of

differing levels of virulence. Virol J 11:46–59

Muggeo V. M. R. 2008. segmented: an R Package to Fit Regression Models with Broken-Line Relationships. R News, 8/1, 20-25. URL <http://cran.r-project.org/doc/Rnews/>.

Nazir J, Spengler M, Marschang RE (2012) Environmental persistence of amphibian and reptilian ranaviruses. Dis Aquat Org 98:177-184. <https://doi.org/10.3354/dao02443>

Northeast Partners in Amphibian and Reptile Conservation. 2015. Disinfection of Field Equipment to Minimize Risk of Spread of Chytridiomycosis and Ranavirus. http://www.northeastparc.org/products/pdfs/NEPARC\_Pub\_2014-02\_Disinfection\_Protocol.pdf

Paetow LJ, Pauli BD, McLaughlin JD et al (2011) First detection of ranavirus in *Lithobates pipiens* in Quebec. Herpetol Rev 42:211–214

Price SJ, Garner TWJ, Cunningham AA, Langton TES, Nichols RA. 2016 Reconstructing the emergence of a lethal infectious disease of wildlife supports a key role for spread through translocations by humans. Proc. R. Soc. B 283: 20160952. <http://dx.doi.org/10.1098/rspb.2016.0952>

Robertson L.S. and R.S. Cornman. 2014. Transcriptome resources for the frogs *Lithobates clamitans* and *Pseudacris regilla*, emphasizing antimicrobial peptides and conserved loci for phylogenetics. Mol Ecol Resour. **14**:178-83. doi: 10.1111/1755-0998.12164.

Simpson, E.H., 1949, Measurement of diversity. Nature, v. 163, p. 688. (Measurement and error bars for dominance, which is the complement of evenness.)

van Regenmortel, M. (1990) Virus species, a much overlooked but essential concept in virus classification. Intervirology 31, 241–254