UNDERSTANDING THE EFFECTS OF HOST AND PATHOGEN DIVERSITY ON DISEASE AND PATHOGEN TRANSMISSION AT MULTIPLE SPATIAL SCALES

by

JOSEPH R. MIHALJEVIC

B.A., Washington University in St. Louis, 2010

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Ecology and Evolutionary Biology

2014

This thesis entitled:
Understanding the effects of host and pathogen diversity on disease and pathogen transmission at multiple spatial scales
written by Joseph R. Mihaljevic
has been approved for the Department of Ecology and Evolutionary Biology

_____
Dr. Pieter T.J. Johnson

_____
Dr. Andrew P. Martin

Date_____

The final copy of this thesis has been examined by the signatories, and we
Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline.

Mihaljevic, Joseph R. (Ph.D., Ecology and Evolutionary Biology)

Understanding the effects of host and pathogen diversity on disease and pathogen transmission at multiple spatial scales

Thesis directed by Associate Professor Pieter TJ Johnson

In wildlife communities, the diversity of both host species and pathogens can affect disease and transmission dynamics. However, the various mechanisms leading to these biodiversity effects occur at strikingly different spatial scales. For my dissertation, I used empirical and theoretical tools to understand how pathogen and host diversity affect transmission at multiple scales. First, I conducted a series of laboratory experiments using a genus of viruses, *Ranavirus*, which can cause devastating die-offs in amphibian populations. I asked how multiple virus types might interact to affect individual-level probabilities of infection and, subsequently, population-level transmission dynamics. I found that co-exposure to two *Ranavirus* species substantially increased the probability of an amphibian larva becoming infected, as well as the average viral load among individuals. Concordantly, the presence of multiple *Ranavirus* species led to larger epidemics in experimental populations, as well as an increased probability of mortality. This research illustrates that *Ranavirus* coinfection could strongly mediate epidemic dynamics in natural amphibian populations. In the next part of my dissertation, I created an epidemiological model in which a single pathogen circulates through a vertebrate host community. I found that the relationship between host species richness and pathogen transmission could be positive, negative, or non-monotonic depending on how the host's total community density scales with host richness and the type of pathogen transmission assumed.

These results highlight that host community composition influences transmission in complex ways, suggesting that observing a consistent effect of host diversity in natural systems is unlikely. Finally, scaling up and using a metacommunity framework, I developed a statistical method to explore how symbiont (including pathogen) communities are structured across space. I then applied this method to a large scale, longitudinal data set of amphibian symbiont communities and discovered that the structure of these communities changes through time and is predominantly influenced by temporal changes in host community composition. Overall, my research illustrates that transmission dynamics are influenced by factors at multiple spatial scales and that integrating across scales is important for understanding how, where, and when biodiversity will affect disease dynamics.

CONTENTS

INTRODUCTION

Wildlife disease is considered a major economic, public health and conservation concern (Daszak 2000, Smith et al. 2006, McCallum 2012). Understanding how pathogens, hosts and environmental context interact to affect disease outcomes and pathogen transmission is an imperative goal and is central to the field of disease ecology (Hudson et al. 2002, Ostfeld et al. 2008). Advances in theoretical modeling and empirical research on host-pathogen interactions have enhanced our ability to explain and, in some instances, successfully manage diseases in both human and wildlife systems (Anderson et al. 1992, Dieckmann et al. 2002, Joseph et al. 2013a). However, while studying single pathogens in single host species has resulted in fundamental insights about epidemiology, evidence is building that community interactions among multiple host and pathogen species can significantly alter expectations of disease outcome and pathogen spread (Rigaud et al. 2010, Tompkins et al. 2010).

One critical finding of disease ecology is that the diversity and structure of host communities can mediate multi-host pathogen transmission and disease severity (Ostfeld and Keesing 2012, Johnson et al. 2013). Emerging field and laboratory data have demonstrated support for the dilution effect, where high plant and wildlife diversity reduces disease severity or pathogen spread in a variety of multi-host pathogen systems (Ostfeld and Keesing 2000, 2012, Keesing et al. 2006, 2010, Johnson et al. 2013b). However, predicting the generality of host diversity-disease relationships in natural systems is difficult. Given the complexity of host-pathogen interactions in a wildlife community context, some researchers have suggested that the dilution effect should not be as general as has been proposed based on current evidence (Randolph and Dobson 2012, Wood and Lafferty 2012). Furthermore, very few studies have

considered how the interactions within parasite communities that are embedded in host communities might further alter disease dynamics.

In natural systems, infection of individual hosts with multiple strains and species of pathogens, known as coinfection, seems to be the rule, rather than the exception (Cox 2001, Telfer et al. 2010, Rigaud et al. 2010, Balmer and Tanner 2011). Recently, it has been proposed that incorporating how multiple pathogens interact within and among hosts is essential to our understanding and management of disease in humans and wildlife (Lafferty 2010, Fenton and Perkins 2010). For example, independent of host community diversity, coinfection can alter expectations of host mortality, pathogen transmission, and even pathogen evolution, often in ways that are context- and system-dependent (Jolles et al. 2008a, Johnson and Hoverman 2012b, Alizon et al. 2013). However it is unclear whether pathogen diversity should scale with disease in the same ways as host diversity, and under which conditions pathogen or host diversity are better predictors of disease patterns.

Understanding how pathogen diversity and host diversity affect disease and transmission patterns deals with questions of scale. Pathogen diversity affects dynamics within hosts, but effects can scale up to the host community-scale. Similarly, host diversity affects transmission at the local community level, but through host dispersal, could affect dynamics at the metacommunity scale. In my dissertation, I use theoretical modeling and a variety of empirical tools to explore this important issue of scale. I have organized my chapters in a way that reflects a spatial hierarchy of questions: from effects of pathogen diversity within individual host individuals and within a host population, to the effects of multi-host species composition on pathogen transmission within local communities, to a series of chapters on exploring the effects of host composition on symbiont metacommunity structure across space and time.

In Chapter 1 I use controlled laboratory experiments with a tractable amphibian-virus system to understand the consequences of viral diversity (via co-exposures) on disease and transmission dynamics in amphibian host individuals and in host populations. In Chapter 2, scaling up to the local host community level, I use multi-species epidemiological modeling and simulation to study, in a generalized framework, how specific aspects of host community composition affect pathogen transmission. In Chapter 3, I conceptually address the ecological and evolutionary significance of embedding symbionts – including parasitic, mutualistic and commensal organisms – and hosts in a metacommunity framework. Then, in Chapter 4 I introduce a novel statistical method that integrates multi-species occupancy modeling in the analysis of metacommunity structure. This method corrects for detection error, which is particularly problematic in studies of symbionts. Finally, in Chapter 5, I utilize these developed methods to analyze a large data set of amphibian symbionts collected over 4 years. I explore how host community composition and exogenous environmental covariates influence symbiont metacommunity structure across space and time.

CHAPTER 1

CO-EXPOSURE TO *AMBYSTOMA TIGRINUM VIRUS* (ATV) FACILITATES GREATER INFECTIVITY AND REPLICATION OF *FROG VIRUS 3* (FV3) IN LARVAL FROGS

1.1 Abstract

Coinfection – the simultaneous infection of a host with multiple pathogen types – is a pervasive natural phenomenon, yet only recently have researchers started to document how it affects disease and transmission dynamics across a variety of systems. In this study, we use two experiments to understand how the individual, host-scale effects of coinfection can scale up to affect population-scale transmission dynamics. Specifically, we co-expose larval frogs to two species of *Ranavirus*, a globally distributed genus of viruses that can cause serious disease and local population die-offs of amphibians, reptiles, and fish. In our first experiment, we expose individual *Rana aurora* larvae to *Ambystoma tigrinum virus* (ATV), wild-type *Frog virus 3* (FV3), or a unique strain of *Frog virus 3* isolated from a ranaculture facility in Georgia, USA, (R-FV3), either in isolation or in pairwise co-exposures. With the aid of viral DNA sequencing, we found that co-exposure to ATV substantially enhanced the infectivity of FV3 and R-FV3, resulting in almost twice as many infected individuals. Co-exposure also increased the average viral loads of FV3-infected larvae. However this same effect was not seen when larvae were co-exposed to the two strains of FV3, suggesting an important effect of viral species identity. In a follow-up experiment, we asked how co-occurrence of ATV and FV3 affects transmission and average viral load in replicate populations of *Pseudacris triseriata*. Although ATV failed to establish in most replicate populations, exposure to ATV resulted in a large epidemic and generally more variable epidemics of FV3. Additionally, individuals in populations co-exposed to both viruses suffered a higher probability of mortality if they became infected. We

hypothesize that these effects of co-exposure are due to immune trade-offs in the face of two unique viruses infecting the host, and emphasize the potentially important role of co-exposure and coinfection by multiple pathogens, even in cases where secondary infections fail to establish. Given the capacity for *Ranavirus* species to cause epidemics and die-offs in amphibians, these findings have important implications for both natural and commercially maintained populations.

1.2 Introduction

Although wildlife disease is now considered a major conservation concern, the disease dynamics of many animal-pathogen systems remain under-studied (Daszak 2000, Smith et al. 2006, Thompson et al. 2010, McCallum 2012). Most ecological studies of disease have historically focused on single host - single pathogen systems, which ignores the reality of natural disease dynamics in many systems (Tompkins et al. 2010). For instance, the simultaneous infection of a single host with a diverse array of pathogen species and strains (i.e. coinfection) is probably the norm in both human and wildlife systems, rather than an exception to the rule (Petney and Andrews 1998, Pedersen and Fenton 2006, Telfer et al. 2010, Balmer and Tanner 2011, Knowles et al. 2013, Griffiths et al. 2014, Nunn et al. 2014). Thus, understanding the consequences of coinfection is an important step forward in the fields of disease ecology and epidemiology (Rigaud et al. 2010).

Coinfection is capable of mediating both the outcome of disease and large-scale disease distribution patterns. Various studies show that infection with multiple pathogen types can result in competitive dynamics within hosts that affect pathogen replication and infectivity (de Roode et al. 2005, Telfer et al. 2010, Johnson and Hoverman 2012a, Johnson et al. 2013a, Nunn et al. 2014) and even pathogen evolution (Choisy and de Roode 2010, Alizon et al. 2011, 2013).

However, fewer studies have attempted to link these individual host-level effects to population-level processes, asking how coinfection can alter epidemiological patterns of transmission and occurrence. For instance, experiments in wild buffalo populations show that coinfection with nematodes strongly influences large-scale patterns of bacterial tuberculosis infection (Jolles et al. 2008a, Ezenwa et al. 2010). In order to understand how coinfection might broadly affect disease and transmission dynamics, more studies are needed to integrate across processes at different spatial scales (Lafferty 2010, Rigaud et al. 2010).

Viruses of the genus *Ranavirus* provide a tractable and relevant model system for exploring the dynamics of coinfection at both the within-host and population-level scales for several reasons. First, ranaviruses infect amphibian communities globally and can cause up to 100% die-off events, constituting a real threat to wild and commercially maintained amphibian populations (Gray et al. 2009b, Lesbarrères et al. 2012). Thus, understanding how coinfection might influence the spread and severity of disease caused by these pathogens is of great concern. Second, this genus of virus is genetically and ecologically diverse, suggesting that interactions among ranaviral species and strains could affect transmission patterns. The two type species of the genus – *Ambystoma tigrinum virus* (ATV) and *Frog virus 3* (FV3) – can be differentiated based on genomic characteristics and their predominant associations with salamanders and frogs, respectively (Chinchar et al. 2009, 2011). Furthermore, many unique strains of both ATV and FV3 differ in key epidemiological traits, such as virulence (Brunner and Collins 2009, Hoverman et al. 2010). Finally, both ATV and FV3 can be highly prevalent across the landscape, and their spatial distributions are broad and likely overlap (Gray et al. 2007, Ridenhour and Storfer 2008, Greer et al. 2009, Brunner et al. 2011, Hoverman et al. 2012), suggesting a high potential for coinfection and possible interactions both within and among hosts. Thus far, however, no study

has considered the role of multiple *Ranavirus* species or strains in mediating disease or epidemics in this system.

While ATV has high infectivity in salamanders (Picco et al. 2007, Brunner and Collins 2009), there is mixed evidence that strains of ATV are able to infect frog larvae (Jancovich et al. 2001, Schock et al. 2008). For example, of the three frog species experimentally exposed to ATV by Schock et al. (2008), all three species showed susceptibility to ATV infection, and a small proportion of individuals died of ATV-induced disease. However, Jancovich et al. (2001) exposed two anuran species to ATV – including a different population of one species also studied by Schock et al. (2008) – and found no signs of infection using both PCR and cell culture isolation. Together, these data suggest that ATV infections in frogs are possible, though the probability of infection likely varies among anuran species and populations and possibly among viral strains. Contrastingly, FV3 infectivity, while varying among species, is consistently high and often leads to mortality in both frog and salamander larvae (Brunner et al. 2005, Picco et al. 2007, Schock et al. 2008, Hoverman et al. 2010, 2011). Given that viral species and strains can cross such host taxanomic boundaries, coinfection is likely a common and important mediator of transmission dynamics in this system.

Theory suggests that the impact of coinfection on transmission dynamics can ultimately depend on its effects on disease and pathogen replication at the host-level scale (Pedersen and Fenton 2006, Jolles et al. 2008b, Mideo et al. 2008, Ezenwa and Jolles 2011a). For instance, if coinfection enhances disease and consequently leads to more rapid host death, then transmission of either coinfecting pathogen could be hindered, leading to smaller epidemics due to coinfection. However, if coinfection facilitates the invasion and within-host replication of pathogens, for example, via immune trade-offs, transmission could be enhanced, leading to

larger epidemics (Ezenwa et al. 2010, Ezenwa and Jolles 2011b). Understanding how coinfection will ultimately affect transmission thus requires multi-scale experiments that determine how effects at the host-level scale up to affect transmission at the population-level.

In this article, we report the results of two experiments conducted to determine the effects of co-exposure to ATV and FV3 in larval amphibians. In the first experiment, we assessed the individual-level effects of co-exposure by exposing individual *Rana aurora* tadpoles to pairwise combinations of ATV, FV3, and a novel strain of FV3 isolated from a ranaculture facility in Georgia, R-FV3 (Miller and Rajeev 2007, Hoverman et al. 2010). However, to also account for potential influences of dosage, we included single pathogen treatments with single and double dose exposures. Our design thus allowed us to determine the effects of co-exposure to two species and to two strains of *Ranavirus*. We monitored mortality daily and assessed ranaviral prevalence and viral load for each treatment using quantitative polymerase chain reaction. In a follow-up experiment, we asked whether the occurrence of ATV and FV3 in microcosm populations of *Psuedacris triseriata* tadpoles affected population-level epidemiological statistics, again assessing mortality, prevalence, and viral load in each microcosm population. For both experiments, we used Sanger sequencing to verify the identity of the dominant virus infecting each individual, which allowed us to narrow down the possible mechanisms leading to an effect of co-exposure.

1.3 Materials and Methods

1.3.1 Viruses and culturing

Aliquots of ATV (Regina ranavirus (RRV) #11800) and FV3 (#061405) were generously provided by Gregory Chinchar. The RRV strain of ATV was originally isolated in 1997 from

Ambystoma tigrinum in Regina, Saskatchewan, Canada (Bollinger et al. 1999), and the FV3 strain is also a wild-type strain isolated from *Rana pipiens* populations of the Midwestern United States in the 1960's (Granoff et al. 1965). An aliquot of the R-FV3 strain of FV3, isolated from a ranaculture facility in Georgia in 2006, was generously provided by Matthew Gray and Debra Miller (GenBank accession no. EF101698; Miller and Rajeev 2007). These viruses were passaged through immortalized fathead minnow cells fed with Eagle's minimum essential medium (MEM) with Hank's salts, containing 5% fetal calf serum. Titer of the resulting viral stocks was determined by plaque assays using serial dilutions of the stock, resulting in titers represented in plaque forming units (PFU). It is important to note that we were unable to obtain an accurate titer of the R-FV3 stock before the start of the first experiment, which likely explains the observed lower-than-expected infectivity.

1.3.2 Experiment 1: Individual-level

This experiment was designed to assess the individual-level effects of co-exposure to ATV and FV3. Egg masses of *Rana aurora* were field collected from wetlands in Oregon in the late spring of 2012 and shipped to the University of Colorado at Boulder, where all experiments took place. Egg masses were first washed with sterile deionized water to remove any possible residual virions and then placed into plastic containers for rearing. *R. aurora* larvae were reared at 20°C with a 12:12 hour day:night photoperiod and fed ground TetraMin® fish flakes (Tetra) *ad libitum* until reaching Gosner stage 30 (Gosner 1960). At this time, larvae were randomly placed into individual, covered plastic containers (with drilled air holes) filled with 1L of carbon-filtered, UV-sterilized water and allowed to acclimatize for 24h. A subset of 15 larvae were humanely euthanized by immersion in an overdose of 1% buffered MS-222 solution and tested

for infection to verify that none of the larvae harbored latent infections prior to experimentation (see quantitative PCR methods below). None of these individuals tested positive for ranaviruses.

Twenty-five larvae were assigned to each of 10 experimental treatment groups: the pairwise combinations of the three viruses (i.e. a single dose of each of two viruses; n=3 treatments), a single and double dose exposure for each virus (n=6 treatments), and a tenth sham control treatment consisting of exposure to a 60µL aliquot of virus-free MEM. On 22 May 2012, a single dose (~$1x10^6$ PFU) or double dose (~$2x10^6$ PFU) of the respective virus or viruses was added to each larva's container via sterile pipette tip. Thus, larvae were passively exposed to each virus inoculate, which likely better mimics field conditions relative to injection-based methods.

After virus was added to each container, individuals were fed *ad libitum* every other day for the extent of the experiment. 100% sterile water changes occurred every 4 days post-exposure to ensure adequate water quality for the larvae. Every measure to avoid cross contamination between containers was used. Standard protocols involved sterilizing dip nets with a 10% bleach solution for 10 minutes, followed by rinsing with sterile water to remove any residual bleach. Container and experimental room surfaces were cleaned with a 2% solution of Nolvasan between each container's water changes, allowed to sit for 10 minutes, and then rinsed with sterile water.

The experiment ran for 21d and mortality of larvae was monitored daily. If an individual died, the individual was extracted from its container, rinsed thoroughly with de-ionized water to remove any non-infecting virions that may have adhered to the individual's skin, and then the entire individual was placed into a microcentrifuge tube and stored at -20°C for later processing. After 21d, all surviving larvae were humanely euthanized by immersion in an overdose of 1%

buffered MS-222 solution. These individuals were then washed thoroughly with de-ionized water, placed into individual microcentrifuge tubes, and stored at -20°C for later processing.

1.3.2 Experiment 2: Population-level.

This follow-up experiment was designed to understand how co-occurrence of ATV and FV3 in a larval amphibian population could alter population-level disease and transmission metrics, which remains relatively rare in coinfection research. Because the first, individual-level experiment showed qualitatively similar effects of co-exposure in the ATV+FV3 and ATV+R-FV3 treatments (see Results below), only ATV and FV3 were used for this experiment. In spring of 2013, we collected egg masses of *Psuedacris triseriata* from local sites in Colorado, washed them with sterile deionized water to remove any possible residual virions, and reared them in plastic containers at 20°C with a 12:12 hour day:night photoperiod. Hatching larvae were fed ground TetraMin® fish flakes (Tetra) *ad libitum* until reaching Gosner stage 30 (Gosner 1960).

The overall design of the experiment was to establish replicate populations of 10 uninfected larvae and then introduce 2 previously virus-exposed larvae each population in order to track the spread of virus and determine if co-occurrence of ATV and FV3 alters the rate of spread and overall epidemic size. We believe that this method of adding infected individuals better mimics natural transmission mechanisms than experimental injection or passive exposure to a large quantity of stock-produced virus.

We randomly assigned a subset of the larvae to the pre-exposed group, which were first anesthetized via immersion in a 1% buffered MS-222 solution. Then these larvae were housed in three batches: FV3-exposure, ATV-exposure, and sham-exposure. Larvae were housed in covered plastic tubs (with drilled air holes) at densities no greater than 1 larva/L water. On 20

June 2013, larvae were passively batch-exposed to a dosage of $5x10^6$ PFU/L water of the respective virus or a sham exposure with an equivalent volume of virus-free MEM and were held in these containers for 4d to initiate infection.

Uninfected experimental populations were established on 20 June 2013 by randomly placing 10 unexposed larvae in 15L covered plastic tubs (with drilled air holes) filled with 12L carbon-filtered, UV-sterilized water. After the 4d batch-exposure, on 24 June 2013, each uninfected population received one of the following combination of exposed larvae: (1) two sham-exposed larvae, (2) two FV3-exposed larvae, (3) two ATV-exposed larvae, or (4) one FV3-exposed larvae and one ATV-exposed larvae. Thus each microcosm population contained 12 total *P. triseriata* larvae (10 sentinels and 2 exposed) for a total density of 1 larva/L water. Each of the four treatments was established with 6 replicate populations.

We had 4 virus treatments (two single-virus exposures, one two-virus exposure, and one sham exposure), two sampling time points (4dpe and 21dpe), and 3 replicate populations at each time point, for a total of 24 microcosm populations of 12 larvae each. After 4dpe a random 3 replicates of each treatment was destructively sampled in order to establish an early epidemic time-point for comparison to late-stage epidemics. Larvae were extracted from each tub, individually, and humanely euthanized by immersion in an overdose of 1% buffered MS-222 solution. As above, euthanized larvae were rinsed, placed into individual microcentrifuge tubes, and stored at -20°C for later processing. Starting at 5dpe, 80% water changes were implemented every 4 days for each remaining replicate population. Mortality was continually monitored, and any deceased individuals were extracted from tubs, rinsed, and stored, as above. At 21dpe, the remaining replicate populations were destructively sampled.

1.3.4 Tissue processing and DNA extraction

Individually stored tadpole carcasses were allowed to thaw to room temperature. 500µL of sterile MEM was added to each carcass, and whole carcasses were then manually homogenized using a motorized homogenizer. This tissue homogenate was then centrifuged at 3000g for 1min. A 500µL aliquot of the resulting supernatant solution was placed into a new sterile microcentrifuge tube to be used for DNA extraction. Qiagen™ DNeasy Blood and Tissue extraction kits and standard protocols were used to extract 250µL of buffered DNA suspension from each supernatant aliquot. DNA samples were stored at -20°C for later processing.

1.3.5 Quantitative PCR amplification of viral DNA

The viral load of each DNA extract (in viral copy number equivalents) was evaluated using quantitative polymerase chain reaction (qPCR), estimated by comparison to a dilution series of standard DNA. We created a synthetic double-stranded DNA standard by synthesizing a 250bp fragment of the major capsid protein (MCP) gene (gBlocks® Gene Fragments; Integrated DNA Technologies™), which is conserved among *Ranavirus* species (e.g. ~97% sequence similarity between ATV and FV3 strains). We used a 10-fold dilution series from $2x10^8$ gene copies down to $2x10^1$ gene copies of standard DNA. Standards and samples were run in duplicate.

The qPCR protocol amplifies a ~70bp region of the MCP, allowing the protocol to identify many *Ranavirus* species; however, consequently, the protocol cannot distinguish between species (Forson and Storfer 2006, Picco et al. 2007). To test each sample, a 2.5µL volume of sample DNA was added to a reaction volume of 17.5µL containing the following reagents: 10µL TaqMan® 2X Universal PCR Master Mix (No AmpErase UNG), 0.06µL forward

primer (for a final concentration of 0.1μM; 5' ACA CCA CCG CCC AAA AGT AC 3'), 0.18μL reverse primer (for a final concentration of 0.1μM; 5' CCG TTC ATG ATG CGG ATA ATG 3'), 0.05μL fluorescent TaqMan® probe (with a starting concentration of 100pmol/μL; 5' FAM-CCT CAT CGT TCT GGC CAT CAA CCA C-TAM 3'), and 7.21μL molecular grade water (Forson and Storfer 2006, Picco et al. 2007). All custom primers and probes were ordered through Life Technologies™. Samples were run in 96-well plates on an Applied Biosciences® machine for 40 cycles: 95°C denaturing (20s), 54°C annealing (20s), and 72°C extension (30s). Two positive ATV and FV3 controls and two negative controls were run on each plate.

After qPCR analysis, the starting sample DNA concentrations of all virus-positive samples was estimated using a Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies™). Thus, all viral loads were standardized to viral copy number per ng of sample DNA. We also quantified the DNA concentration of a random subset of non-infected samples in order to verify that viral detection was not dependent on a high concentration of initial sample DNA.

1.3.6 Viral DNA sequencing of infected samples

We attempted to sequence a small region of the viral genome from all infected samples in order to verify the identity of the infecting virus (or the most dominant infecting virus in each sample). This sequencing should help narrow down the mechanisms driving the effects of co-exposure. We amplified a ~350bp fragment of the MCP gene using a hemi-nested PCR protocol (Kattenbelt et al. 2000). The amplicon from each infected sample, along with a custom sequencing primer (5' ACT ATG CCA CCT CCA TC 3'), was sent to Quintara Biosciences™ for Sanger sequencing. We also amplified and sequenced the same MCP gene fragment from the

three ranaviral strains used in the study (ATV, FV3, and R-FV3). We compared the sequencing data from each infected sample that of the original viral strains.

1.3.7 Statistical Analyses: Experiment 1

All statistical analyses were conducted in the open-source software, R (R Core Development Team 2013). From the first experiment, we had three types of data for each of the 10 viral exposure treatments: survival, proportion of individuals becoming infected, and average viral load per larva. Our overall strategy to compare these metrics consisted of determining if single- and double-dose treatments differed for the single-virus exposures, and then comparing the double-dose single-virus treatments to the three co-exposure treatments and the control. In order to compare survival rates among single- and double-dose single-virus treatments, we conducted a survival analysis (Mantel-Haenszel test), using the '*survival'* package, with dosage as the predictor variable. Because we found no difference in the survival rate between dosages ($\chi_1^2 = 2.6$; P=0.11), we then compared the double-dose single-virus treatments to the co-exposure treatments and control again using the Mantel-Haenszel test. A Cox proportional hazards model yielded the same qualitative results.

We used bias-reduced logistic regression in the '*brglm*' package (Firth 1993) to test the effect of virus species/strain on the proportion of individuals becoming infected in each treatment. First we compared single- and double-dose single-virus treatments by modeling infection prevalence predicted by virus identity (with three levels: ATV, FV3, R-FV3) and dosage. Because dosage did not have an overall effect in the single-virus treatments (z = -0.83, P=0.41), we then compared the control, double-dose single-virus, and co-exposure treatments. For this comparison, we had a fixed effect for each virus type, including all two-way

interactions. Significant interactions would be indicative of non-additive effects of co-exposure.

Finally, to compare viral loads among treatments, we first constructed a linear model predicting the natural log-transformed average viral copy number per ng DNA for each individual (averaged over the duplicate qPCR runs) by dosage. We ran a separate analysis with viral load explained by the day that an individual died, using only the individuals that died. Neither day of death nor dosage correlated with viral load (t = 0.26, P = 0.80; *t* = -1.95, P = 0.07, respectively), so we then ran a similar model comparing double-dose single virus treatments to the co-exposures. We excluded controls from this analysis as none became infected. We ran a model explaining viral load with an interaction between treatment and an additional fixed binomial variable for whether individuals died or not. Again, we ran a separate model with viral load explained by treatment and the day that an individual died, using only the individuals that died. We then ran Tukey's Honest Significant Differences test to compare viral loads pairwise by treatment.

1.3.8 Statistical Analyses: Experiment 2

Similarly to the first experiment, the second experiment had three response variables: survival rate, infection prevalence, and average viral load. However, this time, the response variables were population-specific, with 3 replicate populations per time-point (4dpe and 21dpe) per treatment (FV3, ATV, FV3+ATV, Control).

In order to compare survival among treatments we used a Cox proportional hazards model with replicate population as a random effect (i.e. frailty). Because no control individuals died, only 4 total individuals died in the ATV-only treatment replicates, and no individuals died in any of the 4dpe replicates, we only compared the 21dpe FV3-only and FV3+ATV treatment

replicates to test for an effect of co-exposure.

We compared infection prevalence between the FV3 and FV3+ATV treatments by creating a generalized linear mixed effects model with prevalence explained by treatment, time (as a factor level variable), and their interaction. Replicate population was added as a random intercept term. In order to compare the strength of the random effect between treatments (i.e. the variability among replicate populations), we created two intercept models, one for the FV3-only treatment and one for the FV3+ATV treatment, with prevalence explained only by the random intercept.

We similarly compared the viral load between the FV3 and FV3+ATV treatments by creating a linear mixed effects model with viral load (transformed as in the first experiment) explained by a treatment by time interaction, a fixed effect for day of death, and a fixed effect for whether the individual died or not. Again, replicate population was added as a random intercept term. For all linear models, we performed model simplification by dropping insignificant terms and comparing models with log likelihood ratio tests and AIC values.

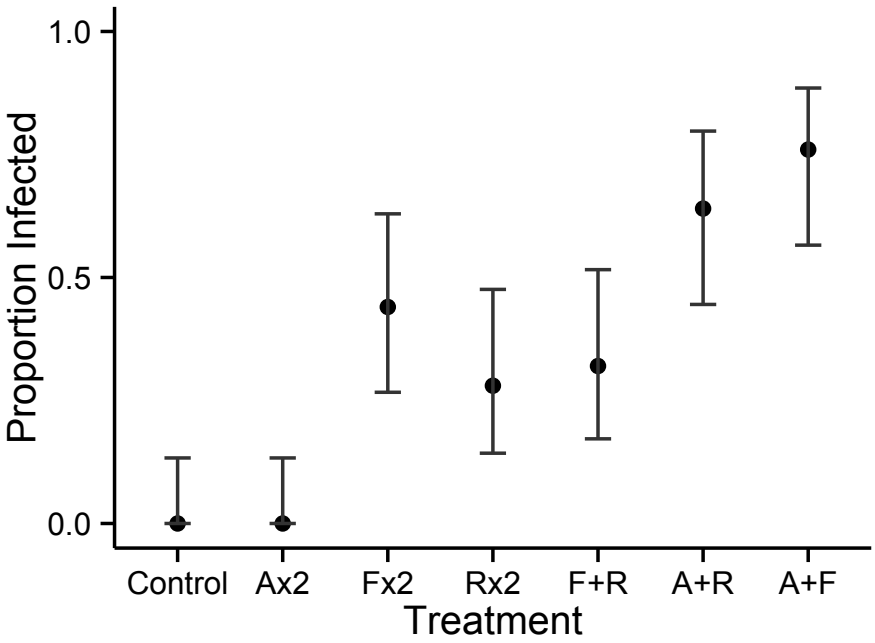
1.4 Results

1.4.1 Experiment 1

While no individuals became infected in the ATV single-virus exposure treatments, co-exposure to ATV and FV3 or RFV3 caused a synergistic effect, significantly enhancing overall infectivity compared to single-virus exposures (Figure 1.1). This effect was supported by a significant negative interaction between FV3 and R-FV3 (z = -3.29, P = 0.001). This interaction was due to the fact that the FV3 + R-FV3 co-exposure did not cause a similar synergistic effect as the ATV co-exposures (Figure 1.1). Specifically, removing the effect of R-FV3 led to a

significantly higher estimated effect of FV3 (with R-FV3: 1.56; without R-FV3: 5.23) and vice versa. In other words, we saw that co-exposure to two *Ranavirus* species, but not to two strains of the same *Ranavirus* species, caused significant increases in the probability of an individual host ultimately becoming infected.

Figure 1.1. Proportion of individuals infected in the double-dose single-virus and co-exposure groups of experiment 1. Error bars represent 95% binomial confidence intervals.



Co-exposure to ATV and FV3 also increased average viral load in infected individuals (treatment effect: $F_{4,54} = 5.78$, $P < 0.001$; Figure 1.2). Although we did not see an increase in load in individuals in the ATV + R-FV3 co-exposure treatment, this was likely due to an inability to obtain an accurate viral stock titer for R-FV3, leading to under-dosing. This error most likely also explains the unexpectedly low infectivity of R-FV3 single- and double-dose treatments compared to wild-type FV3 (Hoverman et al. 2010). As expected, we also found that

individuals that died during the experiment had, on average, higher viral loads compared to surviving individuals ($F_{1,54}$ = 7.94, P = 0.007; Figure 1.3). Among hosts that died, however, there was not a correlation between the day on which an individual died, and their viral load ($F_{1,14}$ = 0.27, P = 0.61).

Figure 1.2. Boxplots of viral copy number per treatment group. The box represents the inter-quartile range (IQR; between first and third quartiles), and the center line marks the median value. The whiskers extends from the box to the highest or lowest value that is within 1.5 x IQR. Data beyond the end of the whiskers are outliers and plotted as points. Letters distinguish significantly different means based on Tukey's Honest Significant Difference test of pairwise differences.

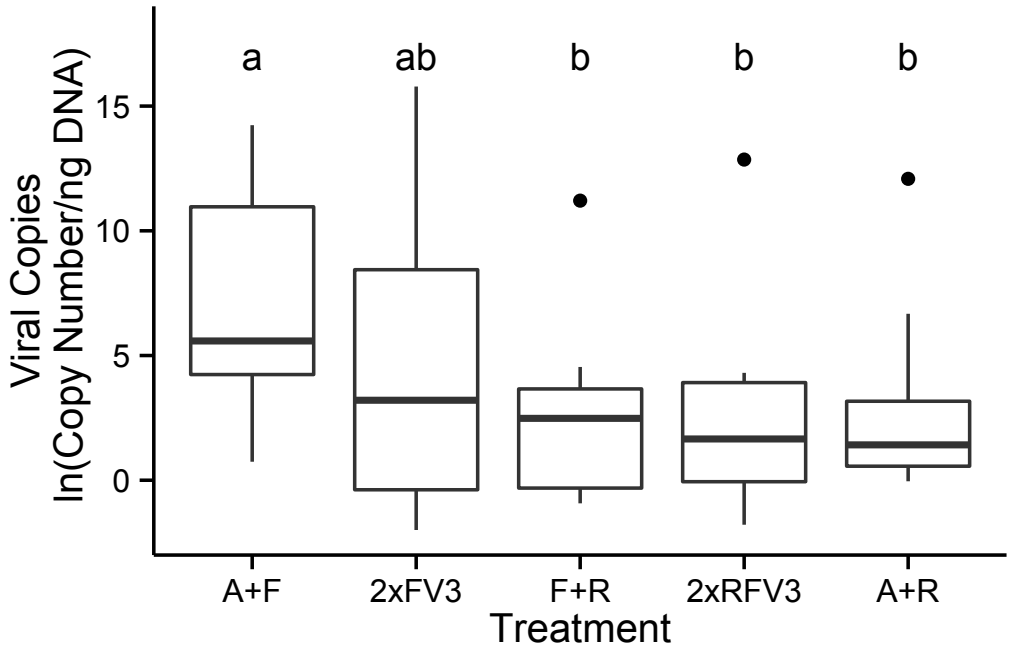
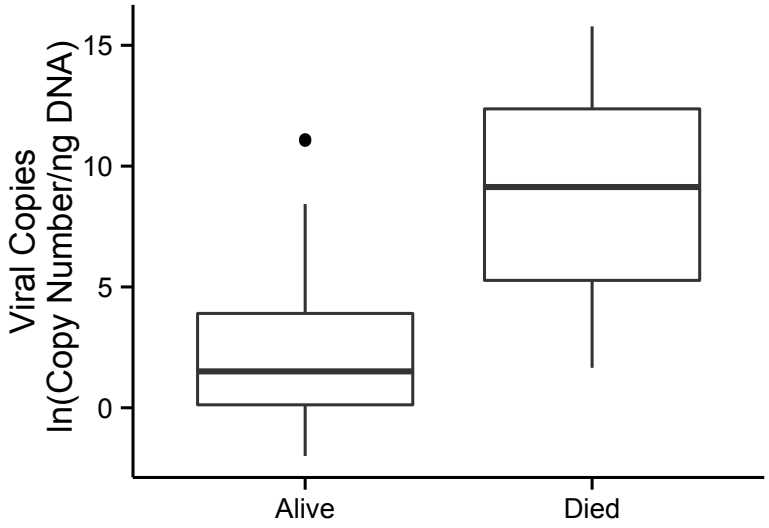


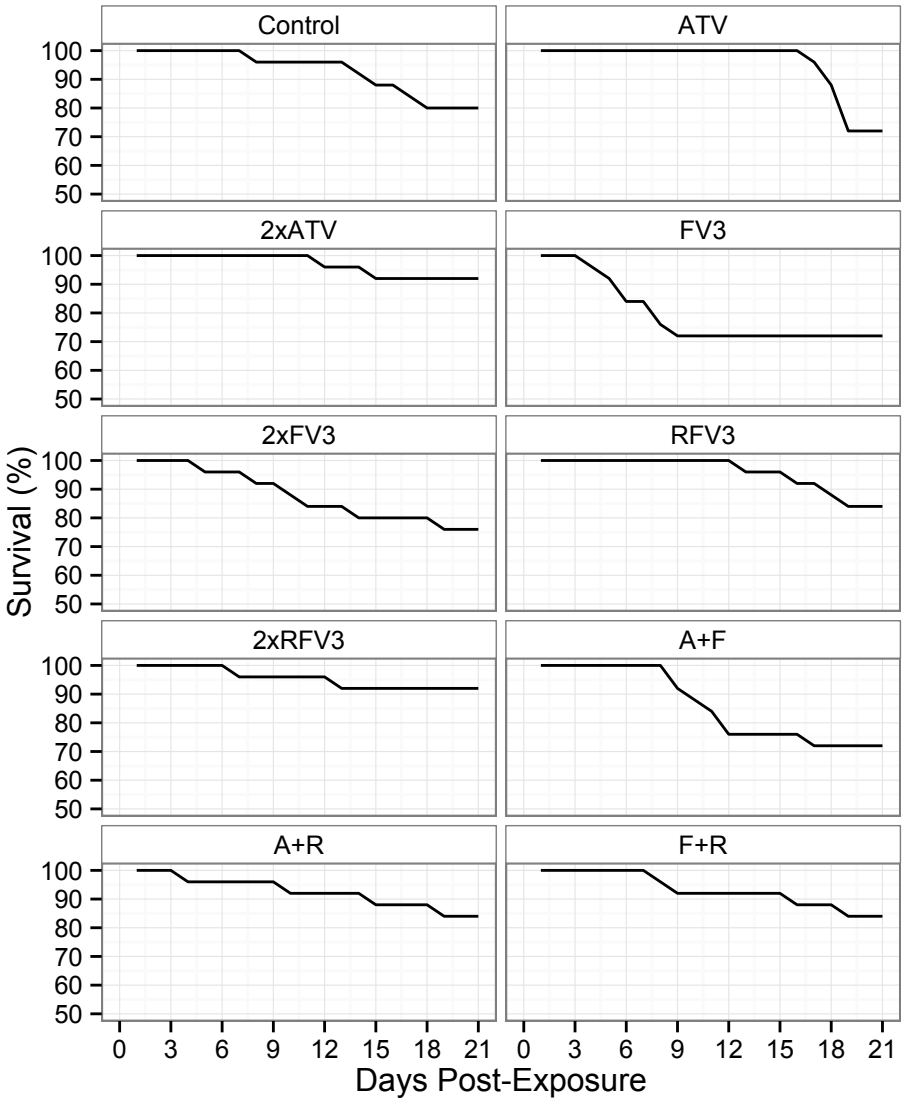
Figure 1.3. Boxplots (as in Figure 1.2) of viral load for infected *R. aurora* that survived until the end of the experiment (Alive) or that died due to infection (Died).



We were able to successfully sequence viral DNA from all but three individuals with low viral load. Sequencing results revealed that all infected individuals were dominantly infected by the FV3-like strains (FV3 and R-FV3) were indistinguishable based on this sequencing method). This result shows that, while ATV was not able to competitively dominate any infections, co-exposure to ATV nonetheless enhanced the infectivity and replication of FV3 compared to the single-exposure treatments. Interestingly, four sequences from the R-FV3 single-virus exposures and the ATV + R-FV3 exposures showed 100% sequence identity to one another, but did not perfectly match the sequences of the three viruses used in this experiment. We searched for similar sequences on GenBank® via BLAST, revealing a 100% match to an isolate of FV3 discovered in lungless salamanders of the Great Smokey National Park, TN (Gray et al. 2009a). The source of this contamination – whether two viruses were co-cultured from the ranaculture facility, or whether the original R-FV3 stock was contaminated post-culturing – is unclear. We do not believe this contamination affected our results.

No individuals from the control treatment were infected during the experiment, although survival did not differ between any double-dose, co-exposure, and control treatments ($\chi^2_6 = 6.2$; P=0.40; Figure 1.4).

Figure 1.4. Survival curves for all treatments in experiment 1.

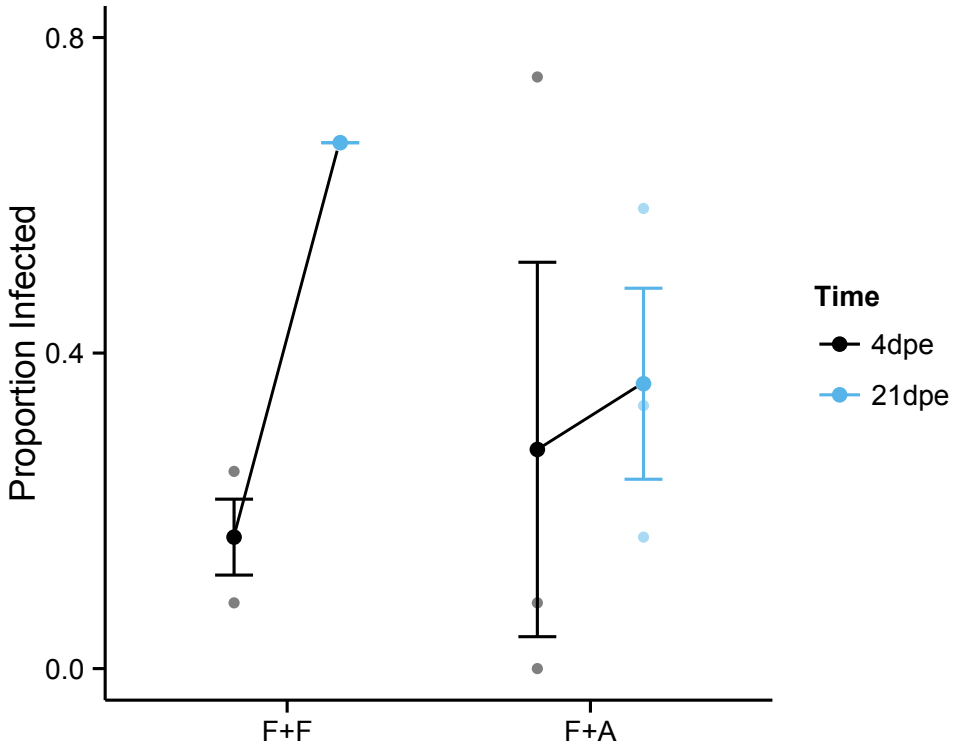


1.4.2 Experiment 2

No individuals in the control replicate populations of *P. triseriata*, and only three individuals in the ATV-only treatments became infected. Therefore, we only used the FV3-only

and FV3+ATV treatment in the analysis comparing infection prevalence. Treatment and time post-exposure interacted significantly to drive prevalence (z = 2.54, P = 0.01); for the FV3-only treatment, the proportion of infected individuals increased more substantially over time compared to the FV3+ATV treatment (Figure 1.5).

Figure 1.5. Proportion of infected individuals between time points and treatments in experiment 2. Time point are distinguished by color, as depicted in the legend. Large, bold circles represent the mean prevalence, and error bars represent one standard error of this mean. Smaller and more opaque circles represent the prevalence of the replicate larval populations. Note that all 3 replicates of the FV3-only treatment at 21dpe had the same prevalence.

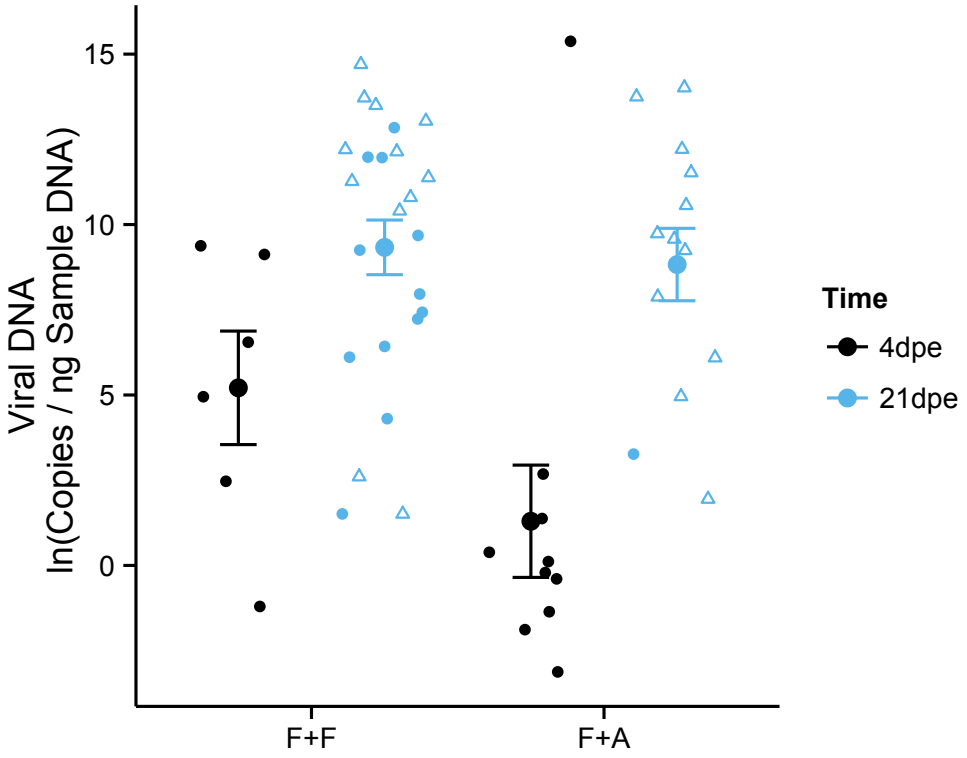


Interestingly, the strength of the random effect of replicates on prevalence was nearly twice as high in the FV3+ATV treatment compared to the FV3-only treatment (variance = 2.08,

1.10 respectively). This was driven, first, by the fact that all three replicates at 21dpe in the FV3-only treatment had the same proportion of individuals infected (8/12). Secondly, in one of the FV3+ATV treatments at 4dpe, 9/12 individuals were infected, which was a substantially larger proportion compared to all other 4dpe replicates, and the highest prevalence in the experiment overall.

In our model for viral load, we only included the data from FV3-only and FV3+ATV replicates. The simplest model explaining viral load among the replicate populations included only time-point, showing that viral load increased over time but was not significantly influenced by whether an individual died or by treatment (AIC=305.83, $t_{4.4} =$ 3.88, P = 0.015; Figure 1.6). However, the model that included a marginally significant effect of whether an individual died and a significant effect of treatment was nearly indistinguishable based on AIC (AIC=308.20; Death: t_{53} = 1.93, P = 0.059; Treatment: t_{53} = -2.06, P = 0.044). In this more complex model, individuals that died had larger loads, and the FV3+ATV replicate individuals had on average smaller loads.

Figure 1.6. Viral copy number across time points and treatments in experiment 2. Time points are distinguished by color, as depicted in the legend. Large, closed circles represent the mean viral load, and error bars represent one standard error of the mean. Smaller closed circles represent the viral load of infected individuals that survived until the end of the experiment (or until destructive sampling in the case of 4dpe replicates). Open triangles represent infected individuals that died prior to the end of the experiment. A jitter is added to the data for ease of interpretation. Notice the most heavily infected individual from the FV3+ATV treatment at 4dpe.



We were able to amplify and Sanger sequence viral DNA from almost all infected individuals, except for 8 individuals with particularly low viral load. Of the three individuals that tested positive for infection in the ATV-only treatments, two DNA samples amplified and sequenced. The sequence data from these two individuals matched that of ATV, showing that *P. triseriata* can become infected with ATV, although rarely. Interestingly, only one individual from all of the FV3-only and FV3+ATV treatment replicates sequenced as being infected with ATV. This individual was present in the replicate population that had both the highest overall

infection prevalence (9/12 in the FV3+ATV, 4dpe treatment), as well as the individual with the largest load of the entire experiment.

When we modeled the death of individuals by viral load and treatment, we found that individuals had a higher probability of death with higher viral loads (z = 2.0, P= 0.039; Figure 1.6) and if they were co-exposed to FV3+ATV (z = 2.1, P= 0.039; Figure 1.6). No individuals from the control treatments, and only four individuals from the ATV-only treatment replicated died during the experiment. Of these four individuals, one tested positive for *Ranavirus*, but unfortunately the viral DNA did not successfully sequence, due to low viral load. A total of 12 and 14 individuals died in the FV3+ATV and FV3-only treatment replicates, respectively. We found no significant difference between the survival in the FV3-only and FV3+ATV treatments ($\chi_1^2 = 0.16$, P=0.68), although there was significant variation among replicates ($\chi_{1.42}^2 = 5.3$ P=0.038).

1.5 Discussion

This study aimed to identify the experimental effects of co-exposure to multiple ranaviruses at the scale of both individual hosts and amphibian host populations. For individual hosts, we discovered that co-exposure to ATV and FV3 increased both the infectivity and within-host viral proliferation of FV3. However, this same effect did not hold for co-exposure to two strains of the same viral species (FV3 and R-FV3), indicating that species identity, but not strain identity is important for predicting the outcome of co-exposure in this system. We then conducted a follow-up experiment at the scale of host populations to understand how these individual-level effects would scale up to affect transmission dynamics. We found limited evidence that, when ATV co-occurs with FV3, co-exposure leads to higher infection prevalence

in the population, as well as more variable epidemic sizes and more variable infection loads among replicate populations. Finally, we found that the probability of death was higher in co-exposure treatments compared to FV3-only treatments. By conducting experiments at both the individual- and population-level scales, we are able to better support that the co-occurrence of *Ranavirus* species has the potential to alter epidemic dynamics in natural amphibian populations and suggest mechanisms for this effect.

Consistent with previous research, ATV was not able to infect larval frogs very effectively in isolation (Jancovich et al. 2001, Schock et al. 2008). In only 4 individuals did we see successful infection, either in *R. aurora* or *P. triseriata*. Importantly, however, co-exposure of ATV and FV3 nonetheless synergistically increased the host's probability of infection. Co-exposure also led to an increase in the average viral load of infected individuals. In our individual-level study, Sanger sequencing revealed that ATV did not dominate any viral infections in the co-exposure treatments. This implies that ATV somehow facilitates the invasion and subsequent proliferation of FV3 within co-exposed larvae. In the second experiment, although ATV only established in one replicate population, this resulted the largest epidemic and an individual with the highest overall viral load, even after only 4dpe. Additionally, in this population-scale experiment, we saw that the probability of mortality given infection was significantly higher for individuals in the co-exposure treatments. 92% of co-exposure infected individuals died, while only 50% of FV3-only infected individuals died by the end of the experiment. Thus, it seems that both of our experiments support an effect of co-exposure on viral infectivity, viral replication, and potentially, mortality.

One hypothesis for a co-exposure-mediated increased infectivity and viral replication in this system is that infection with ATV and FV3 leads to non-overlapping immune responses in

the amphibian larvae. For instance, if the host generates unique adaptive responses to ATV and FV3, then this could lead to a trade-off that decreases the efficacy of the host's response to FV3, facilitating invasion. In our first experiment, this hypothesis seems supported by the fact that co-exposure to two strains of the same viral species (FV3 and R-FV3) did not lead to similar effects, which could be due to an overlapping immune response (i.e. cross immunity) resulting from the DNA sequence similarity of these viral strains.

While there is ample evidence for intraspecific interactions multi-strain infections (Read and Taylor 2001, Mideo et al. 2008, Alizon et al. 2009, 2013), fewer studies have documented the possible immune trade-offs imposed by multi-strain infections (Balmer and Tanner 2011). However, in areas of Senegal with low malaria transmission and consequently low levels of acquired anti-malarial immunity, a larger proportion of cases of clinical malaria are associated with multiple-genotype *Plasmodium falciparum* infections compared to asymptomatic cases (Zwetyenga et al. 1998). Thus, multi-strain infections might pose complex immune challenges that lead to a net decrease in the efficacy of the host's immune response. In the *Ranavirus* system, along with complex innate immune responses, it is known that *Xenopus laevis* adults produce long-lasting anti-FV3 IgY antibodies, and preliminary evidence show similar, though less effective innate and adaptive responses in the larval life stage (Chinchar et al. 2011, Chen and Robert 2011). It is unknown, however if exposure to ATV elicits overlapping innate and adaptive responses with FV3. If the antigens presented by ATV are different enough from FV3, it is possible that the hosts attempt to mount two distinct adaptive responses, decreasing the overall response to FV3 and facilitating invasion and viral proliferation. It will be important for future studies to determine the degree of antibody specificity between ATV and FV3.

With the data from our experiments, we cannot rule out another hypotheses that would

lead to the observed effect of co-exposure. It is possible that if ATV and FV3 coinfect the same host cells, recombination occurs, producing a novel, more infectious virus. Genomic evidence from multiple *Ranavirus* species suggests high recombination frequency and shows that these viruses are prone to host-shifts due to gene acquisition and subsequent adaptation (Jancovich et al. 2003, 2010, Abrams et al. 2013). In fact, recombination has been employed to explain the collinearity and the one inversion between the ATV and FV3 genomes (Eaton et al. 2007). While recombination between ATV and FV3 could occur, it is not immediately evident why such a recombination event would enhance infectivity and viral replication. This is a prime area of future research.

In our second experiment, we also found that the variance in prevalence among replicate populations was substantially larger for the FV3+ATV treatment. This is likely driven by two factors. First, at 4dpe, we saw the replicate population with the largest prevalence overall, driven by the establishment of ATV. Secondly, at 21pde, the FV3-only treatment had no variability in prevalence; each replicate had 8/12 individuals infected. Contrastingly, there was lower overall prevalence and more variability among replicates at 21dpe in the FV3+ATV treatment. This is likely due to an effect of dosage rather than an effect of co-exposure. Because ATV did not establish in the FV3+ATV treatment in any of the 21dpe replicate populations, essentially these populations were only exposed to one infectious individual, whereas in the FV3-only treatments, populations were exposed to two infectious individuals. It is likely that this higher dosage in the FV3-only treatments led to a consistent epidemic peak prevalence (~67%), whereas the lower dosage in FV3+ATV treatment led to more variability in the epidemic peak.

Somewhat surprisingly, we did not see an effect of co-exposure on survival rates in our experiments. Although surprising, this is likely a factor of the longevity of our experiment rather

than a true absence of effect. We ran our experiments for 21d, which in previous studies, has been long enough to see 20-100% mortality due to FV3 infection in other species of frogs and due to ATV in salamanders (Brunner et al. 2005, Hoverman et al. 2010). It is likely that, because co-exposure led to higher viral loads, more individuals would die due to infection if we carried out the experiments for a longer time period.

We suspect that the effect of co-exposure was not as strong in the second experiment because of the difference in viral delivery and dosage. Specifically, ATV was only able to establish in one replicate population (as evidenced by our sequencing methods). This is probably due to the low probability with which our pre-exposed *P. triseriata* individuals ultimately became infected with ATV. In the first experiment, each co-exposed individual was passively exposed to ~$1x10^6$ PFU of ATV, whereas in the second experiment, uninfected sentinels could only become exposed to ATV if the pre-exposed individuals became infectious. Thus, it is likely that if we had passively exposed the replicate populations to ATV in a way similar to our first experiment, a larger effect would be seen. Future studies could reveal such a dosage effect by exploring how the ratio and overall dosage of FV3:ATV influences the likelihood of co-exposure effects at the individual and population scales. The differences in effects between the two experiments could also be due to differences in the effect of co-exposure among amphibian species. It is known that variability in FV3 infectivity among amphibian species has phylogenetic and ecological correlates (Hoverman et al. 2010, 2011). It will be important to determine if ATV + FV3 co-exposure affects certain species more greatly than others.

Our results illustrate that in natural amphibian populations, co-occurrence of ATV and FV3 could substantially alter epidemic cycles. Specifically, if ATV can establish in the population, co-exposure with FV3 could result in more infected individuals and subsequently

higher mortality rates in the long run. This effect seems particularly relevant for wetlands in which salamanders and frogs co-occur. If ATV is present and infects the local salamanders and FV3 establishes in the anuran populations, spillover of ATV from the salamanders could enhance FV3 epidemics in the frogs. Also, because FV3 is adept at infecting salamanders as well (Schock et al. 2008), it is likely that such a scenario would increase infection prevalence and intensity in the urodele population. Thus, our results illustrate the need to consider co-exposure and co-infection in the amphibian-*Ranavirus* system and emphasize the need for field data on ATV and FV3 co-occurrence at both the wetland- and host individual-levels.

CHAPTER 2

THE SCALING OF HOST DENSITY AND RICHNESS CAN CHANGE THE DIRECTION, SHAPE, AND DETECTIBILITY OF DIVERSITY-DISEASE RELATIONSHIPS[1]

2.1 Abstract

Pathogen transmission responds differently to host richness and abundance, two unique components of host diversity. However, the heated debate around whether biodiversity generally increases or decreases disease has not considered the relationships between host richness and abundance that may exist in natural systems. Here we use a multi-species model to study how the scaling of total host community abundance with species richness mediates diversity-disease relationships. For pathogens with density-dependent transmission, non-monotonic trends emerge between pathogen transmission and host richness when host community abundance saturates with richness. Further, host species identity drives high variability in pathogen transmission in depauperate communities, but this effect diminishes as host richness accumulates. Using simulation we show that high variability in low richness communities and the non-monotonic relationship observed with host community saturation may reduce the detectability of trends in empirical data. Our study emphasizes that understanding the patterns and predictability of host community composition and pathogen transmission mode will be crucial for predicting where and when specific diversity-disease relationships should occur in natural systems.

2.2 Introduction

Emerging field and laboratory data lend support to the dilution effect, where high plant and wildlife diversity often reduces disease severity or pathogen spread in a variety of multi-host

[1] Published as: Mihaljevic, JR, MB Joseph, SA Orlofske and SH Paull. (2014). The scaling of host density and richness can change the direction, shape, and detectability of diversity-disease relationships. *PLoS ONE* 9(5): e97812.

pathogen systems (Ostfeld and Keesing 2000, 2012, Keesing et al. 2006, 2010, Johnson et al. 2013b). Although multiple mechanisms should lead to a dilution effect, many others could underlie the opposing pattern, termed the amplification effect (Keesing et al. 2006). To date, predicting the generality of diversity-disease patterns in natural systems has proven difficult, and given the complexity of host-pathogen interactions, some have suggested that the dilution effect may be less common than previously expected (Randolph and Dobson 2012, Wood and Lafferty 2012).

The specific mechanisms driving diversity-disease patterns have been debated extensively in the literature, particularly because various ecological and epidemiological properties of host communities can influence the spread of pathogens. For example, both host richness and host abundance (or density) are expected to affect diversity-disease trends (Dobson 2004). A dilution effect is expected if species rich communities have more host species that are resistant to infection, demonstrating a role of richness *per se* in limiting disease. For instance, Johnson et al. (Johnson et al. 2008) experimentally controlled host abundance, finding a direct effect of larval amphibian richness on reducing trematode infection in American toad (*Bufo americanus*) larvae. However, a dilution effect can also be seen if the abundance of a species that strongly contributes to pathogen reproduction and transmission (i.e. a highly competent focal host) negatively correlates with host richness. Notably, Mitchell and colleagues (Mitchell et al. 2002) found that, of 11 foliar fungal pathogens of plants, roughly half showed a dilution effect due to reduced focal host abundance, rather than a richness effect.

Given that both host richness and host abundance affect pathogen transmission, the relationship between richness and total community abundance should affect how pathogen transmission scales with host richness. Theoretical exploration of this topic has considered two

species accumulation types: (1) additive, where total community abundance scales linearly with richness, and (2) compensatory, where total community abundance is invariant to species richness (Dobson 2004, Rudolf and Antonovics 2005). These extreme scenarios generate unique null expectations of how transmission should scale with diversity (Rudolf and Antonovics 2005). However, a more realistic expectation might be that community abundance saturates with increasing richness. For example, saturation of total community biomass and percent cover has been documented in various plant systems (Tilman et al. 1996, 2001, Guo et al. 2006). Despite these observations, the effects of saturating host abundance on pathogen transmission have not been explored theoretically or empirically.

Saturating host abundance could lead to the intermediate result between completely additive and completely compensatory richness-abundance relationships, that being a non-monotonic trend between richness and pathogen transmission. Recently, researchers have resorted to *reductio ad absurdum* reasoning to argue that non-monotonic trends between species richness and disease risk must occur in disease systems with observed dilution effects (Wood and Lafferty 2012, Lafferty and Wood 2013). Specifically, if zero host species were present at a site, there would be no disease, so that adding any number of susceptible hosts species would initially increase disease. Then, as richness increases, an inflection point (i.e. dilution effect) may be observed (Wood and Lafferty 2012, Lafferty and Wood 2013). However a quantitative exploration of potential non-monotonic diversity-disease trends grounded in community ecological theory relevant to many disease systems is still lacking.

Here, we build upon previous models to explore how varying the empirical relationship between total host community abundance and host richness affects community-level disease patterns. First, we consider the effects of additive, compensatory and saturating host abundance

on pathogen transmission in simulated multi-host communities under both density- and frequency-dependent transmission scenarios. We predict that a more realistic saturating abundance-richness relationship will reveal more complex patterns, including non-monotonic relationships between host richness and pathogen transmission. Using simulation, we also investigate the effect of various abundance-richness relationships on the detectability of diversity-disease patterns. We find that non-monotonic relationships between host richness and pathogen transmission can occur under certain conditions, but that high variability could lead to low detectability of such trends.

2.3 Methods

The mathematical model used here is a multi-species extension of the classic susceptible-infected-recovered (SIR) epidemiological model (e.g. Dobson 2004):

$$dS_i/dt = b_i N_i - d_i S_i - S_i \sum_{j=1,n} \beta_{ij} I_j \ ,$$

$$dI_i/dt = S_i \sum_{j=1,n} \beta_{ij} I_j - \left(d_i + \alpha_i + \sigma_i\right) I_i \ ,$$

$$dR_i/dt = \sigma_i I_i - d_i R_i \ .$$

We assume that all host individuals are susceptible, infectious or recovered (immune for life), designated with S, I and R, respectively. All other parameters are defined in Table 2.1, and below we describe how values for each parameter are assigned to different host species.

Table 2.1. Parameter assignment and definitions for creating the species pool and epidemiological model.

Parameter	Value	Definition	Biological Explanation
Construcing Global Species Pool (Preston's Law)			
P_i	1 - 8	Preston's rank	The rank for each species, which corresponds to the assigned abundance
z	0.1	Constant derived from field data	Scales the difference in abundance from one rank to the next, with the modal rank as reference
M	3	Modal rank	The mode of the distribution of abundances among all species in the sample community
Y_o	10	Number of species present in the modal rank.	Species richness in the modal abundance rank
z	0.1	Constant derived from field data	Scales the difference in abundance from one rank to the next, with the modal rank as reference
K_i	2 - 256	Abundance at each rank, assigned on a $\log_2$ scale	The abundance at carrying capacity of a particular species in the community
s	1 - 10	Number of species in each rank	The outcome of Preston's law, which determines how many species have a given equilibrial abundance
Assigning species traits			
w_i	$\log(w_i)=a-b\log(P_i)$	Species specific weight	Weight is determined by rank-abundance, so that more abundant species are smaller
a	2	Constant	Scales the relationship between species abundance and body weight
b	1	Constant	Scales the relationship with species abundance and body weight
R_{0i}	0 - 2	Intraspecific reproductive number of the pathogen for each host species	Determined by a truncated gamma distribution, such that most species are poor hosts ($R_{0i}<1$). More abundant, and therefore smaller, species are assigned higher R_{0i} values
k	0.3	Constant	Determines the scale of the gamma distribution from which R_{0i} is drawn
ψ	3	Constant	Determines the shape of the gamma distribution from which R_{0i} is drawn
Epidemiological model			
b_i	$0.6w_i^{-0.27}$	Birth rate	Species birth rate determined by allometric scaling with body size
d_i	b_i	Death rate	Species death rate assumed to be equal to birth rate
α_i	$(m-1)d_i$	Pathogen induced mortality	Decrease in mean lifespan due to infection, proportional to death rate. Scales with body size so that larger species have lower pathogen induced mortality

m	1.5	Constant	Determines the proportionality between species death rate and pathogen induced mortality rate
σ_i	εd_i	Recovery rate	Species ability to recover from, and become immune to, infection
ε	10	Constant	Determines the proportionality between life span and recovery
β_{ii}	$R_{0i}\,(d_i+\alpha_i+\sigma_i)/K_i$	Per capita, intraspecific tranmission rate under density-dependent transmission	The ability of an infected individual in the community to contact and successfully transmit the pathogen to another individual of the same species under the assumption of density-dependent transmission
β_{ij}	$c_{ij}(\beta_{ii}+\beta_{jj}/2)$	Interspecific transmission rate	The ability of an infected individual of one species in the community to contact and successfully transmit the pathogen to another individual of a different species
c_{ij}	0.05	Constant	Scaling parameter controlling the amount of intra- and interspecific transmission among species in the community

2.3.1 Constructing a global host species pool

In many multi-host pathogen systems, host species vary in ecological and epidemiological traits relevant to pathogen transmission. For example, in the Lyme disease system, mammal hosts ranging from mice to raccoons to deer can all become infected with the bacterial pathogen and spread this pathogen to tick vectors. To summarize a very complex system, each host species varies in its population dynamics and in its ability to acquire and transmit the pathogen to ticks; therefore, the community composition of hosts is very important for determining overall pathogen transmission, and subsequently, disease risk (LoGiudice et al. 2003). In our model, we attempt to construct a global host species pool that captures the ecological and epidemiological variability seen in generalist pathogen systems. Thus, we draw heavily upon established trends in community ecology and allometric scaling laws to assign host species plausible parameter values. We derived our epidemiological model and allometric scaling laws from Dobson (Dobson 2004), and integrated Roche et al.'s (Roche et al. 2012,

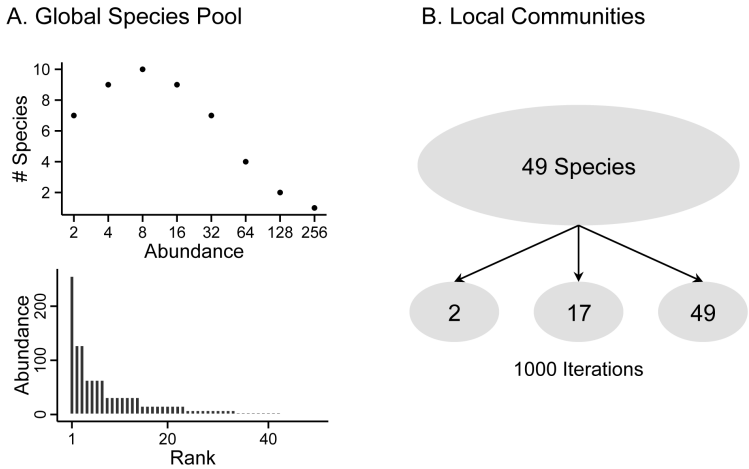
2013) methods of generating realistic communities using Preston's law. In contrast to previous models, this work focuses primarily on assessing the consequences of more realistic host richness-abundance scaling on diversity-disease relationships.

We assembled a global pool of vertebrate host species following Preston's law of abundance distributions. This law generates a lognormal distribution of species' abundances, where most species are rare and only a few species are abundant, a pattern observed in many natural communities (Preston 1948, Roche et al. 2012). Preston's law is given by:

$$s(P) = Y_0 e^{-(zP-M)^2},$$

where z is a constant, s is the number of species that are present in the Pth rank from the modal rank, M, and Y_0 is the number of species that are present in the modal rank (Figure 2.1; Table 2.1). This creates a Gaussian-type curve that dictates how many species have certain population sizes (Figure 2.1A). Thus, species were assigned equilibrium abundances, K_i, according to their given rank. For all analyses, our global pool consisted of the same 49 host species.

Figure 2.1. Conceptual diagram of assembling the global host community, species traits, and local communities. *A*, Preston's octaves of abundances and resulting rank-abundance of the 49 host species used in our model. *B,* Schematic of our methods for choosing 1000 local communities. Species in local communities were chosen at random.



Each species' weight, w_i, was obtained from a scaling relationship (Table 2.1). We assume body size scales exponentially with abundance, so that the most common species were the smallest. Birth rates, b_i, were derived from allometric scaling, so that larger species reproduce less frequently (Leo and Dobson 1996, Roche et al. 2012). We additionally assumed that death rates, d_i, were equal to birth rates, a common assumption of populations at equilibrium. Pathogen induced mortality, α_i, was assumed to be a proportional decrease in mean lifespan due to infection (Dobson 2004). This means that larger species, which have lower background mortality rates, also have lower death rates due to infection. Recovery rates, σ_i, were assumed to scale with death rates, so that larger species have slower recovery rates compared to smaller species (Table 2.1).

Intraspecific R_0, R_{0i}, values describe the pathogen's growth rate in each host species' population, by taking into account intraspecific transmission and the average duration of infection:

$$R_{0i} = \frac{\beta_{ii} K_i}{\left(\alpha_i + \sigma_i + d_i\right)}$$

in the case of density-dependent transmission. Similar to Roche et al.'s (Roche et al. 2012) treatment of 'susceptibility', we drew realistic species-specific R_{0i} values from a right-skewed truncated gamma distribution, ranging from zero to two, where an $R_{0i} \geq 1$ means the pathogen can invade that species' population. This gamma distribution resulted in most R_{0i} values being close to or less than 1, so that the pathogen could not invade most host species' populations in isolation, but a few host species could sustain (small) epizootics. For example, white-footed mice are very competent hosts for the pathogen that causes Lyme disease, but most other mammal hosts (e.g. squirrels) are much less competent (LoGiudice et al. 2003).

Additionally, many bird species can harbor West Nile Virus, but only a few of these species are responsible for passing infections on to mosquito vectors (Kilpatrick et al. 2006).

We derived intraspecific transmission rates β_{ii} corresponding to intraspecific R_0 values for each species (Table 2.1). Host competence in our model is thus defined as the probability of transmission given a contact between a susceptible and infectious individual, regardless of species identity, which is proportional to β_{ii} and R_{0i}. We further assumed that R_{0i} was negatively correlated with body size and positively correlated with abundance. Therefore, the smallest, most abundant host species were the most competent hosts [e.g. 18,21]. This pattern might be expected if pathogens have selective pressure to adapt to more common species, and/or if larger host species (which in our model have lower population sizes) have selective pressures to invest more heavily in pathogen defense strategies in order to survive to reproductive age. For example, in the Lyme disease and West Nile systems, host body size correlates negatively with host competence (Huang et al. 2013). Furthermore, there is evidence that more 'fast-lived' and abundant amphibian species experience higher infection intensities and more severe pathology from trematode parasites (Johnson et al. 2012, 2013b). A recent review of the literature suggests that there is evidence that life history traits, such as body size and abundance, are correlated with host competency; however, the strength of these correlations are often unclear, and this variability across more disease systems needs to be further assessed (Joseph et al. 2013b).

Finally, interspecific transmission rates, β_{ij}, were calculated as the pair-wise average of intraspecific transmission rates of species, i and j. The strength of interspecific transmission was controlled by a scaling parameter, c_{ij}, in the form:

$$\beta_{ij} = c_{ij}\left(\frac{\beta_{ii} + \beta_{jj}}{2}\right).$$

Because intraspecific transmission rates, β_{ii}, vary across host species, there is some inherent

heterogeneity in interspecific transmission in our global communities, which is scaled with the c_{ij} term.

2.3.2 Simulating local communities

We considered three different conditions governing the relationship between species richness and community abundance, as well as density-dependent and frequency-dependent transmission for each of the three conditions. The first condition, termed the "additive" method, assumes species abundances in simulated communities are equal to their abundance in the global pool, which leads to a positive linear relationship between species richness and community abundance. The second condition, termed the "compensatory" method, fixes community abundance regardless of species richness, but the abundance of all species is proportional to their relative abundance in the global pool. In other words, if a species is common, its abundance is adjusted to still be common with respect to the other species in the community. These first two conditions correspond to completely additive and compensatory abundance assumptions, respectively, investigated by Rudolf and Antonovics (Rudolf and Antonovics 2005), but generalized to the *N* species case. Importantly, the "compensatory" method is also analogous to experimental designs that vary host richness but fix total density of hosts to isolate the effect of richness. The third assumption, termed the "saturating" method, imposes a curvilinear relationship between community abundance and species richness. Because of a lack of empirical data informing the nature of such a relationship in vertebrate communities, we use two different curvilinear relationships: asymptotic and logistic curves (Appendix).

To investigate the relationship between community composition and pathogen transmission, we iteratively simulated local communities by drawing random subsets of species

from the global pool (Figure 2.1B). Species richness in each local community thus ranged from 2 to 49 species. We simulated 1000 local communities for each set of conditions described above.

Under each scenario described above, we calculated community R_0, a measure of potential pathogen transmission in a naïve local host community (Dobson 2004, Allen et al. 2012). This metric is analogous to the population-level R_0 but is extended to incorporate interspecific transmission. When community $R_0 \geq 1$, the pathogen can invade and persist in the host community. Values above 1 correspond to larger epizootic sizes, as community R_0 also correlates with maximal infection prevalence in the community. We also calculated the coefficient of variation of community R_0 for each value of host richness in order to assess how the variability of pathogen transmission changes across the range of host richness.

Community R_0 is calculated as the dominant eigenvalue (spectral radius) of the N x N matrix (G) that incorporates the rate of transmission between species and the average duration of infection for an individual of the species transmitting the infection, based on the SIR model:

$$G = \begin{pmatrix} \frac{\beta_{ii} p_{ii}}{(\alpha_i + \sigma_i + d_i)} & \cdots & \frac{\beta_{ji} p_{ji}}{(\alpha_i + \sigma_i + d_i)} \\ \vdots & \ddots & \vdots \\ \frac{\beta_{ij} p_{ij}}{(\alpha_j + \sigma_j + d_j)} & \cdots & \frac{\beta_{jj} p_{jj}}{(\alpha_j + \sigma_j + d_j)} \end{pmatrix}$$

Here the p terms vary whether transmission is frequency or density-dependent. For density-dependent transmission, p is equal to abundance of the infecting species (rows) at the disease-free equilibrium, K_i. For frequency-dependent transmission, p is equal to the relative abundance of the infecting species at equilibrium, a proxy for the relative proportion of interspecific contacts (i.e. $K_i \Big/ \sum_i^N K_i$) (Dobson 2004, Allen et al. 2012).

Thus, community R_0 is essentially determined by each species' host competence, abundance, and the strength of interspecific transmission. In order to verify that our assumptions about how life-history traits relate to host competence (e.g. positive relationship between abundance and intraspecific R_0) did not strongly affect our results, we also created a 'null' model. For this model, we randomized all life-history traits to eliminate associations with intraspecific R_0, R_{0i}. We then derived intraspecific transmission rates β_{ii} to match R_{0i} for each species and used this community in the simulations described above.

2.3.3 Simulating the effects of sample size on detecting diversity-disease patterns

In empirical field studies, researchers are limited by sample size and sample breadth (i.e. the number of sites that are available to sample, and the range of host richness observed). To investigate how abundance-richness patterns affect the detectability of diversity-disease relationships with variable sampling effort, we simulated sets of independent communities ranging from sets of 5 communities to sets of 45 communities and calculated community R_0 for each community in each set. For each simulated set of communities (i.e. sets with different sample sizes) we built a general additive model (GAM) of community R_0 predicted by host species richness using a cubic regression spline with shrinkage, using the 'mcgv' package in R (Wood 2011, Team 2013). This modeling approach allows for the detection of curvilinear (including non-monotonic) trends in the data. Due to low sample size in the smaller community sets, we limited the maximal number of knots on the spline to three (Keele 2008). We replicated this method 20 times for each value of sample size, totaling 820 simulations for each scenario (below).

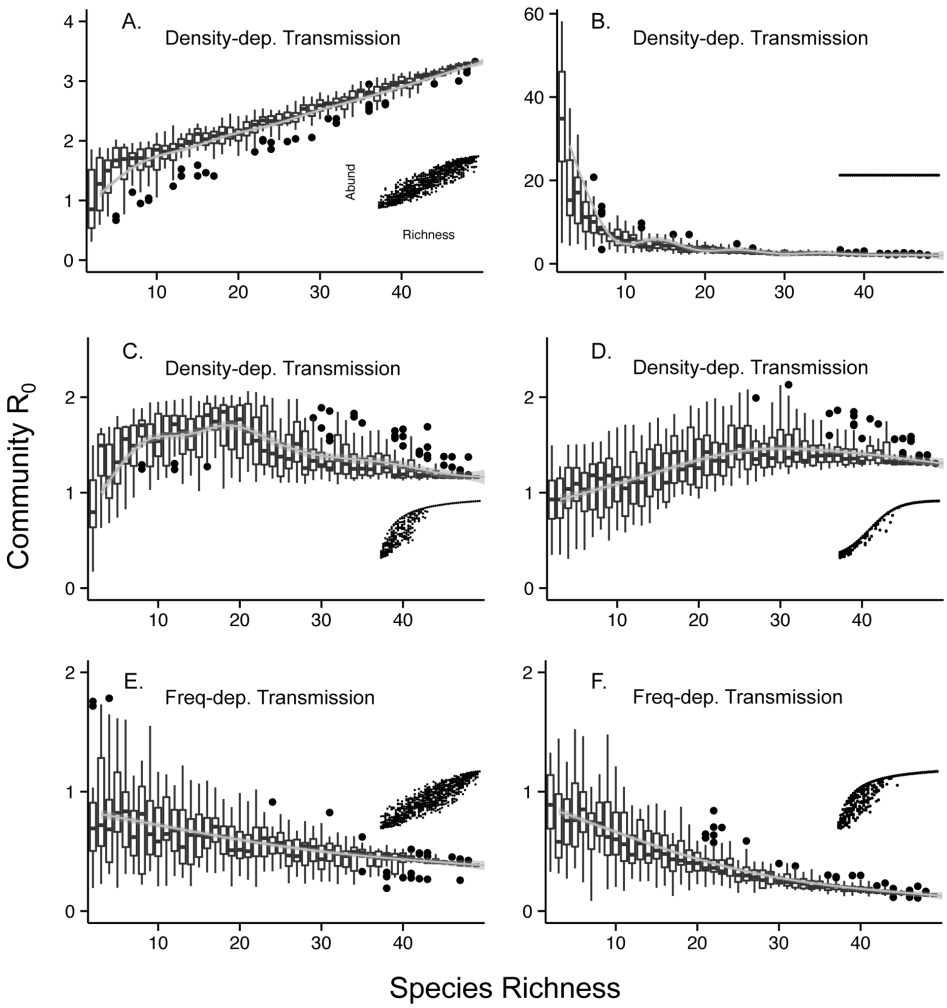
We conducted the GAM on simulations from three different scenarios (treatments): (1)

the “additive” method of simulating abundance-richness relationships under density-dependent transmission, (2) the “additive” method under frequency-dependent transmission, and (3) a “saturating” abundance-richness relationship under density-dependent transmission. We limited this analysis to only three treatments – as opposed to all of the scenarios explored above – for statistical tractability. We used logistic regression to determine how sample size and treatment affected the detectability of significant relationships between community R_0 and host species richness. All simulations and statistical analyses were conducted in R (R Core Team 2013).

2.4 Results

We found that a saturating abundance-richness relationship with density-dependent transmission led to a clear non-monotonic trend in which there was an initial increase in community R_0 (i.e. an amplification effect), followed by a decrease in community R_0 at a higher range of richness values (i.e. a dilution effect; Figure 2.2C-D). The degree to which the pattern was non-monotonic was influenced by the community abundance-richness relationship assumed; nonetheless, the non-monotonic pattern seemed general over a range of values for these assumptions (Appendix). This pattern persisted under our null model, with random associations between host competence and abundance (Appendix), demonstrating that our results are not sensitive to this assumption.

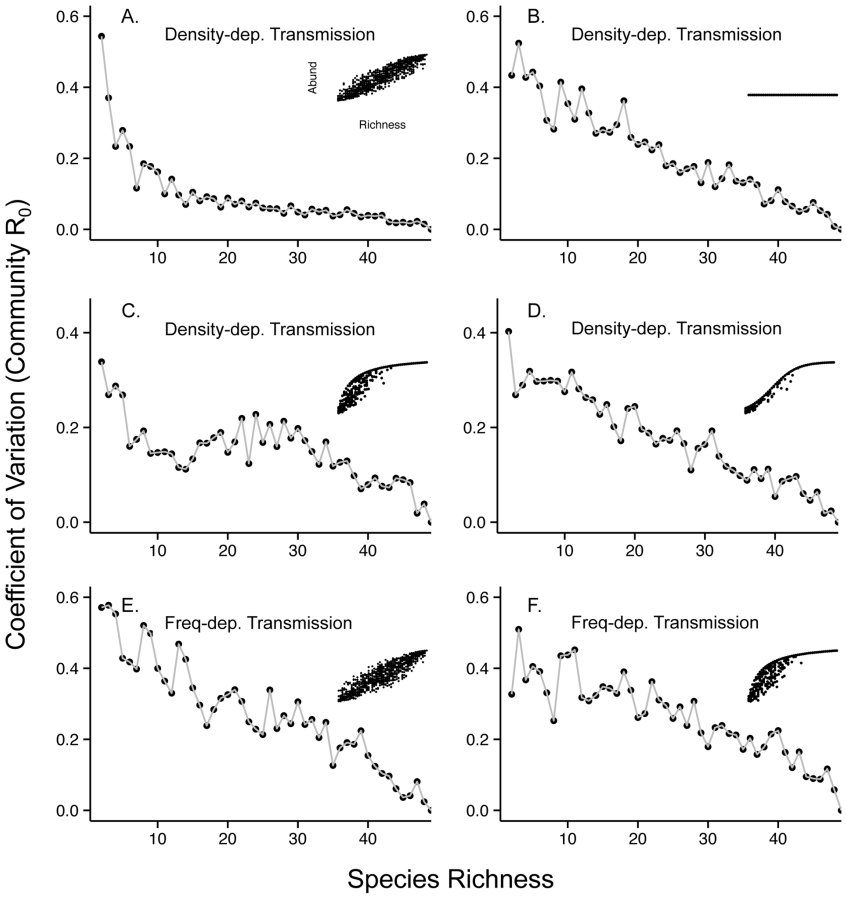
Figure 2.2. The relationship between community R_0 and host species richness for six example scenarios. Panels *A-D* show results from simulations based on the four different assumptions of the underlying relationship between host community abundance and richness (depicted as inset Figures) with density-dependent transmission. Panels *E* and *F* are two examples with frequency-dependent transmission. Boxplots summarize the findings of 1000 simulations for each panel. LOESS smoothers with 95% confidence bands were added for visual interpretation of average trends. Not all iterations of frequency-dependent transmission are shown because they show the same qualitative trends. (Parameters used to generate these data: Y_0=10, z=0.10, M=3, a=2, b=1, m=1.5, ε=10, k=0.3, ψ=3, c_{ij}=0.05).



As expected, with density-dependent transmission and a linear relationship between community abundance and richness (i.e. the "additive" method), community R_0 monotonically increased with host richness (Figure 2.2A). Also as expected, when community density was kept constant across the full range of host richness (i.e. the "compensatory" method), there was a

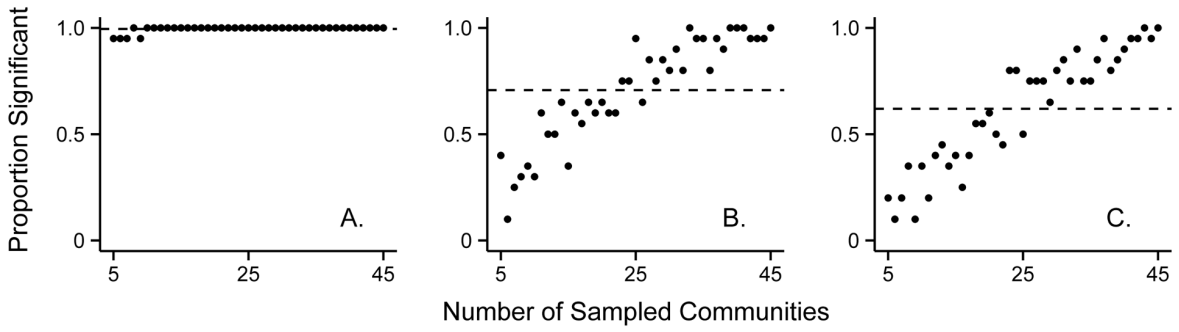
marked decrease in community R_0 as richness increased (Figure 2.2B). This effect was exaggerated under our null model scenario, because in this case some typically rare species were randomly assigned high host competence (Appendix). Finally, under frequency-dependent transmission, community R_0 decreased as richness increased regardless of the relationship between community abundance and host richness assumed (e.g. Figure 2.2E-F). We also found that the coefficient of variation in community R_0 decreased markedly with increasing host richness irrespective of the community abundance-richness relationship and the mode of transmission (Figure 2.3).

Figure 2.3. The coefficient of variation of community R_0 at each value of richness for the simulated communities shown in Figure 2. The underlying relationships between community abundance and richness are shown as inset Figures. Parameters are as in Figure 2.



The high variability in our simulations, as well as the saturating abundance-richness pattern affected the probability of finding a significant relationship between community R_0 and richness with increasing sample size (Figure 2.4). Across all three treatments – (1) "additive" with density-dependent transmission, (2) "additive" with frequency-dependent transmission, and (3) "saturating" abundance-richness relationship with density-dependent transmission – the probability of finding a significant relationship increased markedly with sample size (z=18.37, P<0.0001; Figure 2.4). Furthermore, we were overall less likely to find significant relationships between community R_0 and richness in the case of a saturating abundance-richness relationship, compared to the two additive cases (additive, density-dependent: z=10.56, P<0.0001; additive, frequency-dependent: z=4.37, P<0.0001; Figure 2.4). The main effects of sample size and treatment explained much of the variation in finding significant trends (pseudo-R^2=0.38). We included an interaction between treatment and sample size in the initial model, but this term was insignificant and was dropped from the final model.

Figure 2.4. Results of GAM to test the effect of community abundance-richness relationships and pathogen transmission mode on community R_0-richness relationships across a range of sample sizes. *A*-*C*, Proportion of simulations where the GAM was significant versus sample size, for the three treatments: *A*, "additive" method with density-dependent transmission; *B*, "additive" method with frequency-dependent transmission; and *C*, "saturating" method with density-dependent transmission. The horizontal dashed lines in *A*-*C* show the total proportion of significant cases across all sample sizes (i.e. out of 820 simulations) for each of the three treatments. Parameters of generated local communities follow those specified in Figure 2.



2.5 Discussion

Evaluating the generality of diversity-disease relationships in nature is a difficult task due to the complexity of host-parasite interactions, challenges involved in achieving replication, and extrinsic environmental factors influencing pathogen transmission (Rigaud et al. 2010, Tompkins et al. 2010, Randolph and Dobson 2012). Using a multi-host species epidemiological model, we found that the relationship between total host community abundance and host richness can mediate how pathogen transmission scales with host richness. Particularly, saturating host abundance-richness relationships can lead to situations in which the community R_0 of a pathogen with density-dependent transmission increases over low ranges of host richness but decreases over higher ranges. Moreover, across all abundance-richness patterns and the two pathogen transmission modes explored in this study, the variation observed in community R_0 was much higher in low richness communities compared to speciose communities. We also found that community density saturation may reduce the detectability of statistically significant diversity-disease relationships.

Here we demonstrate that understanding the ecology of fundamental host community dynamics can improve predictions about when and where to expect the dilution effect to occur. Our model results support previous predictions that generalist pathogens with density-dependent transmission are likely to increase in prevalence when species additions are additive and decline when they are compensatory (Dobson 2004, Rudolf and Antonovics 2005). We find that frequency-dependent transmission disease dynamics did not respond to abundance-richness relationships, because transmission is independent of host density. By contrast, for the case of saturating community abundance-richness relationships with density-dependent transmission, we found an amplification effect for the portion of the curve where host communities accumulate

species abundance additively. Then, as more speciose communities start to saturate and transition to compensatory species additions, we observe a dilution effect.

An important assumption in our model is that host competence was strongly, positively correlated with species abundances. While this assumption is pervasive in the diversity-disease literature, the few studies explicitly testing for such a relationship show mixed results (Joseph et al. 2013b). A recent modeling study showed that variability in the strength of the relationship between host competence and extirpation risk (assumed to be correlated with life-history traits such as abundance) can lead to mixed dilution and amplification effects when communities are completely subtractively or substitutively disassembled (Joseph et al. 2013b). Using a complementary null model, our results show that regardless of the assumed relationship between host competence and species abundance (or other life-history traits), a saturating relationship between total community abundance and species richness can result in non-monotonic diversity-disease trends with density-dependent transmission. This observation of non-monotonic diversity-disease trends with saturating abundance-richness relationships is general because of how species abundances accumulate with species richness. At low richness, we are seeing dynamics driven by additive abundance-richness relationships, which then transition to being driven by compensatory relationships as the community saturates. This general pattern persists under the null model due to the strong influence that density has on community R_0 (via contact rates), independent of the relationships among demographic parameters that result from allometry, life history trade-offs, or pathogen adaptation. Therefore, we propose that as long as abundance saturates with richness, and density-dependent transmission occurs, a non-monotonic diversity-disease trend is likely.

A saturating abundance pattern is more likely in vertebrate communities when resources

are abundant with a low number of species, but competitive interactions become more pronounced as species richness increases. For example, using three different theoretical models, Lehman and Tilman (Lehman and Tilman 2000) found that a saturating relationship between total community biomass and richness emerges due to competitive interactions among species. Other abundance-richness relationships not explored in this study could emerge due to particular community dynamics that vary with host richness or are tied to particularly influential species. For instance, predators may influence disease dynamics more strongly than competitors in certain systems, and therefore the shape of richness-abundance relationships may depend on the richness of communities that predators tend to occupy. Additionally, the presence of ecosystem engineers or keystone species may affect the shape of abundance-richness relationship in ways specific to particular study-systems. It will be important for future ecological and disease studies to determine how total host abundance scales with richness as communities both assemble and disassemble, and the predictability of these trends across disease systems, in order to evaluate how often curvilinear or non-monotonic community R_0-richness patterns might occur in nature.

Our model findings also emphasize that the range of host richness values observed in a field study could be important for determining the specific diversity-disease relationship that is detected. For example, Guo et al. (Guo et al. 2006) showed that non-monotonic relationships between biomass production and grassland plant richness exist, but only when a wide enough range of richness is sampled. In our model, if the host community size is smaller than that required for the community to begin saturating (e.g., if host community < 20 species in Figure 2.2D), an amplification effect would be the logical expected outcome of diversity loss, but this expectation would change if more speciose communities were sampled. It should be noted, however, that in our model, the threshold of 20 species before a dilution effect is a product of the

model structure and assumptions. Identifying possible saturation thresholds in natural communities is an important consideration when generating null expectations for how diversity loss contributes to disease risk for pathogens with density-dependent transmission. Future field and laboratory studies could assess these expectations by manipulating host abundance-richness patterns and observe if non-monotonic relationships can arise across a wide range of host richness.

This model also supports the idea that, in some instances, host species identity can be more important for driving diversity-disease relationships than richness *per se* (LoGiudice et al. 2008, Randolph and Dobson 2012). We found that regardless of the abundance-richness relationship and transmission mode assumed, the variability in community R_0 declined markedly with increasing host richness. We can attribute this pattern to a sampling effect that emerged due to the fact that the number of unique communities that could be assembled at low values of richness exceeded those at higher values of richness. This means that, by chance, combinations of species with very high competence or very low competence could be present at low values of richness, suggesting that species identity plays an important role in low-richness communities. This emergent property conforms to various biological traits that show similar declines in variability with increasing richness due to statistical sampling effects, termed the 'variance reduction effect' (Huston 1997). Mitchell and colleagues (Mitchell et al. 2002) found a similar pattern of decreasing variability in fungal pathogen load in plant communities with increasing richness, which they also attribute to stochastic species dominance at low host richness. Therefore, it could be the case that host species identity is more important for determining pathogen transmission in communities of low richness, compared to more speciose communities. This could be especially true in cases where host species with high pathogen competence tend to

occupy species poor communities more frequently than species with low pathogen competence (Johnson et al. 2013b).

Saturating abundance-richness relationships could also obscure diversity-disease patterns in the field due to sampling issues. For example, we found that the non-monotonic community R_0-richness pattern produced by a saturation scenario translated into finding fewer significant regressions across a range of sample sizes, even when using general additive models that can detect such curvilinearity (Figure 2.3). Additionally, the high degree of variation observed in the "additive" case with frequency-dependent transmission resulted in many non-significant regressions, even though the general pattern between community R_0 and host richness was clearly negative (Figure 2.2E). Furthermore, the power to detect diversity-disease relationships would be lower in real studies where metrics of disease or host diversity are estimated, rather than known exactly (as in the case with our model and community R_0).

Our findings also have implications for the experimental design of studies that are investigating diversity-disease relationships. In our simulations, when the total abundance of the host community was kept constant (i.e. completely compensatory additions), a dilution effect is invariably seen, as long as interspecific transmission rates are low (Figure 2.2B). This scenario is analogous to experiments that fix host density or abundance to isolate the effects of host richness. This pattern occurs because as richness increases, each species' density declines, resulting in less within-species transmission. If interspecific transmission is low, then these epizootics do not readily spill over, causing smaller community-wide epizootics. Researchers should measure interspecific transmission rates before designing an experiment that attempts to isolate the effects of host richness on disease trends, especially in systems of generalist pathogens with density-dependent transmission. Experimental designs should also, in as much as

possible, incorporate host abundance and species composition data from the field to more accurately represent natural transmission dynamics in the lab.

Conclusion

Evidence from field, experimental and theoretical studies increasingly suggests that the details of host community composition, and not just species richness *per se*, are important for driving diversity-disease patterns. Previous research has shown that total host richness and the abundance of hosts can moderate disease patterns, and that host species identity can be a more important predictor of disease risk than host richness. Here we have demonstrated that an often-overlooked metric of host community composition – the scaling of total host community abundance with host richness – may drive previously unpredicted non-monotonic richness-pathogen transmission relationships. These non-linear trends, as well as generally high variability in pathogen transmission in depauperate host communities, tend to hinder pattern-detection with low sample sizes.

Our model adds to a growing body of work that suggests that finding generalizable diversity-disease patterns in the field across host-pathogen systems may be more difficult than previously appreciated. However, in our model, high variability in pathogen transmission is often driven by the random nature of local community composition. More data are needed to understand how communities assemble and disassemble in terms of host richness, host abundance and host competence. For instance, it has been proposed that depauperate communities may be primarily inhabited by highly abundant, competent host species (Keesing et al. 2010, Johnson et al. 2013b). Joseph et al. (Joseph et al. 2013b) suggest that it is reasonable to expect that on average more competent hosts occupy species poor communities due to host life history traits and pathogen evolution, but that even slight variability around competence-

extirpation risk relationships can cause mixed disease dilution and amplification throughout community disassembly. Bringing community ecologists and disease ecologists together in order to better understand the predictability of host community composition and host competence - as well as their relative contributions to diversity-disease trends - would greatly aid in building more informative models of when and where diversity should affect disease.

CHAPTER 3

LINKING METACOMMUNITY THEORY AND SYMBIONT EVOLUTIONARY ECOLOGY[2]

3.1 Abstract

Processes that occur both within and between hosts can influence the ecological and evolutionary dynamics of symbionts, a broad term that includes parasitic and disease-causing organisms. Metacommunity theory can integrate these local and regional scale dynamics to explore symbiont community composition patterns across space. In this article I emphasize that symbionts should be incorporated into the metacommunity concept. I highlight the utility of metacommunity theory by discussing practical and general benefits that emerge from considering symbionts in a metacommunity framework. Specifically, investigating the local and regional drivers of symbiont community and metacommunity structure will lead to a more holistic understanding of symbiont ecology and evolution and could reveal novel insights into the roles of symbiont communities in mediating host health.

3.2 Expanding symbiont community ecology

The study of symbionts and symbiont communities, whether these are commensal, mutualistic, or parasitic organisms, is vital to our understanding of general host-symbiont dynamics as well as clinical and epidemiological patterns. Exploring the community-level ecological interactions of symbionts and their hosts has substantially added to our knowledge of host-symbiont relationships (Pedersen and Fenton 2006, Graham 2008, Ezenwa et al. 2010, Telfer et al. 2010) as well as symbiont evolution (Mideo et al. 2008, Mideo 2009, Alizon et al.

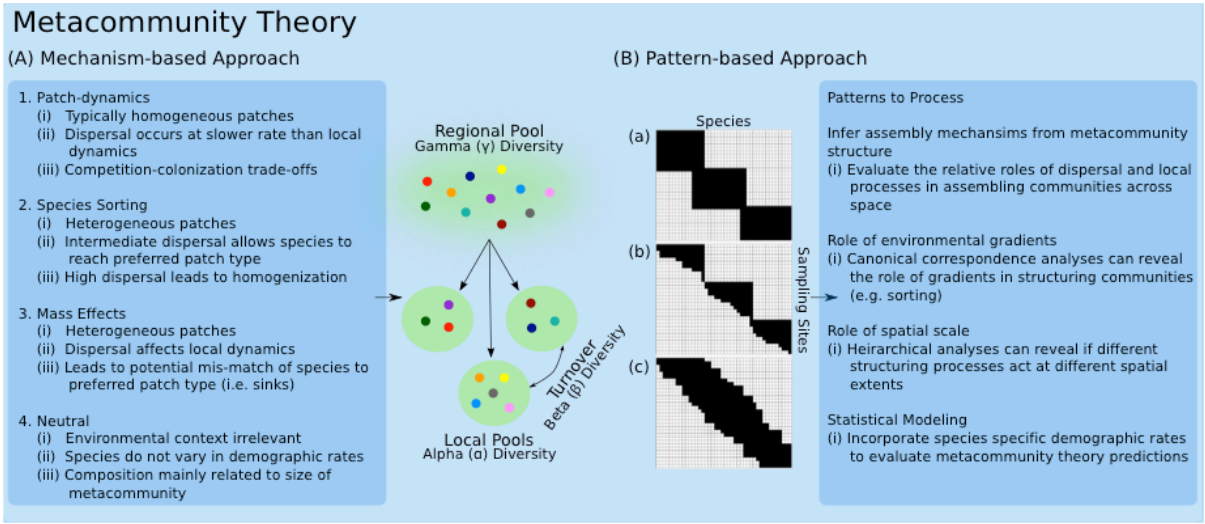
---

[2] Published as: Mihaljevic, JR. (2012) Linking metacommunity theory and symbiont evolutionary ecology. *Trends in Ecology & Evolution* 27:323–329.

2011). Furthermore, symbiont communities themselves have contributed to the fields of ecology and biogeography, for example, by testing theoretical predictions such as distance decay of similarity and niche breadth-range size relationships (Poulin et al. 2011). However, there is an extensive body of other ecological theory that can be applied to and tested with symbiont communities, which will enhance our understanding of host-symbiont interactions, including disease dynamics, and of the structuring of ecological communities in general. Here I advocate the utility of metacommunity theory (Leibold et al. 2004, Holyoak et al. 2005; Figure 1)

The ultimate aim of metacommunity theory is to evaluate patterns of and mechanisms contributing to species diversity across space. Specifically, this theory asks how dispersal of organisms between communities alters local dynamics and subsequently influences community structure both locally and regionally (Box 3.1; Figure 3.1). Furthermore, the emerging 'evolving metacommunity' concept explores how genetic variation is distributed across space and how gene flow can influence species interactions and community composition via (mal)adaptation (Urban and Skelly 2006, Urban et al. 2008, Urban 2011). The metacommunity concept has been integral to our understanding of large-scale trends in community structure and biodiversity (Ricklefs 2008, Cavender-Bares et al. 2009, Leibold 2011). The goals of metacommunity theory and the predictions that have emerged from this body of work align well with current research on symbiont communities and host-symbiont interactions.

Figure 3.1. A summary of metacommunity theory. Metacommunity theory can be generally divided into two approaches: the (A) mechanism-based and the (B) pattern-based approaches. The mechanism-based approach utilizes four modeling paradigms to ask how the regional species pool partitions into local habitats and how these local communities vary across space. The four paradigms differ mainly in the role of patch heterogeneity and the timing and effect of dispersal on local dynamics. The pattern-based approach determines the structures of natural metacommunities. Metacommunities (a-c) are cartoon examples of such structures. These represent presence/absence matrices that are ordinated via a process such as reciprocal averaging. Then other ordination methods and null models are used to associate this structure with a variety of potential biotic and abiotic structuring mechanisms.



Interspecific interactions among symbionts, as well as their dispersal and transmission, all influence the ecological and evolutionary dynamics of symbiont populations and communities, often in complex ways. For example, direct and indirect interactions among multiple symbionts within an individual host can affect the ways symbionts influence host health and can also affect the evolution of symbiont traits, such as the virulence of parasitic symbionts (de Roode et al. 2005, Bell et al. 2006, Choisy and de Roode 2010, López-Villavicencio et al. 2011). These intra-host dynamics are linked with those of larger spatial scales (e.g. inter-host) via symbiont dispersal or transmission. Understanding how these 'local' and 'regional' processes interact is at the forefront of symbiont-related research, especially that of parasitic symbionts

(Mideo et al. 2008, Alizon et al. 2009, Rigaud et al. 2010). Fortunately, merging local and regional scale dynamics is the motivation behind metacommunity theory.

Despite the conceptual similarities with free-living species' metacommunities, the consideration of symbiont metacommunities and their relevance to host health, to symbiont eco-evolution and to the study of ecological communities has not been addressed in depth. Here, I emphasize that considering the metacommunity dynamics of symbionts will reveal a suite of novel and highly relevant questions pertaining to the structuring of ecological communities and to the influence of symbiont communities on host health and functioning. I explain how symbiont metacommunities fit well within the conceptual framework of the standard free-living organism metacommunity, even though clear differences exist. These differences, however, could actually serve to broaden the scope and test the assumptions of metacommunity theory. I then outline the novelty of metacommunity theory compared to other methods of studying symbiont dynamics and discuss various practical and general applications that stem from linking these two fields.

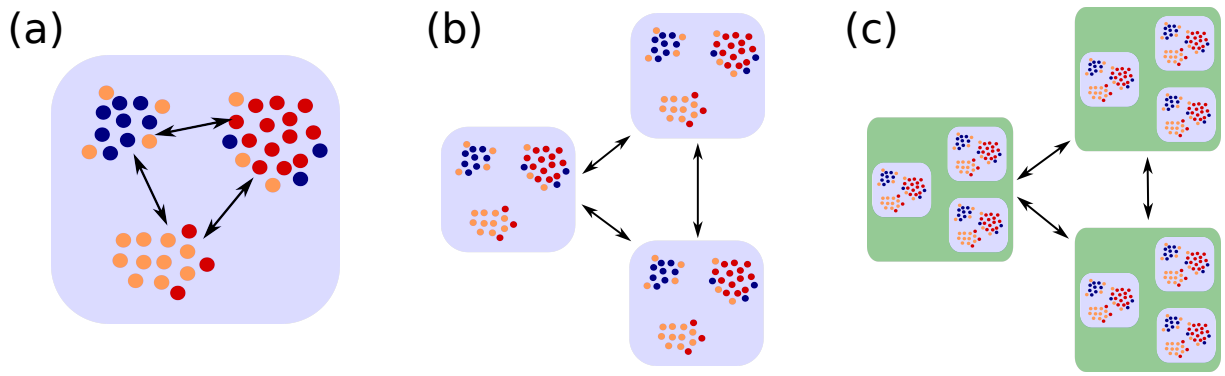
3.3 Defining a (symbiont) metacommunity

Briefly, a metacommunity consists of multiple local communities of interacting species that are connected by the dispersal of at least one of those species. Local communities can be limited to competing 'guilds' of species or can consist of more complex trophic webs. The important aspect here is that dispersal of species among localities changes local community dynamics, leading to community structures that deviate from those expected when considering closed communities. This allows researchers to study community dynamics at larger - and more realistic - spatial scales than previously considered in community ecology (Ricklefs 2008).

Although the spatial delineation of 'local' communities is stringent in theoretical studies, the principles derived from metacommunity theory have been successfully applied to natural communities that lack definitive boundaries (e.g. grasslands, forests; (Holyoak et al. 2005)).

'Local' symbiont communities can be designated at various spatial scales, and, thus, multiple metacommunities can emerge depending on the spatial scale under consideration and the specific questions being addressed (Box 3.2, Figure 3.2). While I believe that each spatial scale is worth considering and can lead to novel research initiatives, for clarity, here I will consider the implications of studying symbiont metacommunities in which a local community consists of the suite of symbionts that inhabit a single host individual (e.g. Figure 3.2b). In this way, the local scale is clearly spatially delineated as a single host in which symbiont population dynamics unfold. Emigration from a host is achieved by the release of infective or dormant symbiont life stages. This numerically changes the demographic parameters of resident symbiont populations within the host, and thus, potentially changes the community interactions among symbionts. Colonization of a new host occurs via dispersal or transmission of symbionts and the subsequent initiation of population dynamics in that new host, which, as above, can affect the local community dynamics.

Figure 3.2. Representations of symbiont metacommunities at various spatial scales. (a) Various host compartments (e.g. organs) house different suites of local symbiont communities that are connected by dispersal of symbionts (represented by arrows). (b) Host individuals serve as local communities or patches for symbionts. Here, patches are connected by dispersal or transmission of symbionts from host to host (represented by arrows). (c) Host sub-populations serve as patches for symbionts, where the suite of symbionts within a host subpopulation constitutes the local symbiont community. In this scenario, patches are connected by the migration of hosts from one sub-population to another, which augments the local suite of symbionts. Different colored dots represent symbiont individuals and species.



There are two clear distinctions between symbiont metacommunities, as defined above, and those of free-living organisms. First, for symbionts, the local habitat patch (host) is not necessarily static in space; however, this is not an insurmountable challenge. For example the metapopulation framework - a single species perspective that follows the same spatial assumptions as the metacommunity framework - has long been advocated as useful for understanding parasite infection dynamics (Grenfell and Harwood 1997, Thrall and Burdon 1997). The disease metapopulation framework has been implemented across spatial scales to explain parasite infection patterns in cases where parasites infect multiple host tissues (i.e. intra-host metapopulations (Frost et al. 2001)), hosts are considered patches (G. R. Hess et al. 2002), or parasites are dispersed among host sub-populations via host migration (Hess 1994, McCallum and Dobson 2002). Furthermore, the theory of island biogeography (MacArthur and Wilson 1967), developed with respect to static islands, has been fruitfully applied to explain prevalence

patterns of parasites when mobile hosts are considered “islands” (Kuris and Blaustein 1980, Reperant 2010).

The success of metapopulation and island biogeography theory in studying parasites makes sense because symbiont population dynamics unfold within a host whether or not the host is mobile. Thus, if hosts were to be considered closed communities, their movement would be irrelevant. The movement of hosts, however, can clearly affect the dispersal and colonization rates of symbionts, especially in the case where direct host-to-host transmission occurs.

Host movement is a unique feature of symbiont metacommunities that can be integrated into metacommunity theory and lead to interesting research questions. First, the timescale and relative role of dispersal in structuring metacommunities varies depending on the theoretical paradigm considered, and, thus, symbiont metacommunities could be analyzed to determine which paradigm is most relevant (Box 3.1). Second, understanding the role of environmental heterogeneity in structuring communities across space is a principal focus of metacommunity theory (Box 3.1). The rate of host movement could then be used as a type of patch heterogeneity, in which symbionts in a more vagile host have distinct demographic rates compared to those residing in less vagile hosts. Finally, while most theoretical models that have formed the foundation of metacommunity theory consider the spatial distribution of patches implicitly, newer methods explicitly model this distribution (e.g. (Pillai et al. 2011)). Thus, a variety of host structures could be simulated in which these ‘patches’ are more or less connected, representing the relative contact rate of hosts.

The second distinction involves the mechanisms that result in patch vacancy. In free-living organism metacommunities patches become fully or partially vacant and can be subsequently recolonized if, for example, residing species go extinct due to demographic

stochasticity or strong interspecific interactions. However, when hosts are considered patches for symbionts, patches can become vacant and subsequently *unavailable* to re-colonization if hosts die due to parasitism or develop strong resistance to parasitic symbionts. Thus, the number of available patches does not remain constant, as is assumed in current implicit-space metacommunity models.

Perhaps the issue above has caused reluctance to use mechanistic metacommunity models to study symbionts, especially because epidemiological models already account for such host-symbiont dynamics. Nevertheless, I feel that, again, this caveat is not intractable. Implicit terms could be added to models for patch destruction (e.g. via pathogen virulence or host resistance), and then balanced with patch creation (e.g. via host immigration or birth rate) (G. R. Hess et al. 2002). In fact, host resistance to symbiont infection could be modeled more realistically as underlying patch heterogeneity. Then, one could explore how intraspecific variation in these host characteristics affects local and regional composition of symbionts (Osnas and Dobson 2011). Finally, using explicit space models, host loss could be modeled similarly to habitat destruction (Pillai et al. 2011). However, the evolution of parasite virulence and host resistance, a common question in epidemiological modeling studies, might be more difficult to address using these methods. The evolving metacommunity perspective, described below, might hold more promise for these important questions. Below I will highlight the novel features of metacommunity theory compared to methods of studying infection dynamics in light of their respective research agendas. I will also more fully develop the practical and general utility of linking metacommunity theory and symbiont evolutionary ecology.

3.4 The utility of considering symbiont metacommunities

3.4.1 The novelty of metacommunity theory

Metacommunity theory differs from current methods of studying symbionts, especially epidemiological models, in both method and agenda. Epidemiological models deal with infectious, parasitic symbionts. These are mechanistic models that follow the numbers or densities of infected, susceptible and resistant hosts in a spatially implicit or explicit (e.g. network analysis) manner, incorporating the negative effects of the pathogen. Epidemiological models ask questions about, for example, the probability of epidemics occurring in the host population and the evolution of pathogen traits, such as virulence, in response to within- or between-host dynamics. Importantly, recent models are able to integrate within-host infection processes and between-host transmission dynamics (Mideo et al. 2008, Day et al. 2011). While some studies have considered the evolutionary consequences of multiple pathogen strains or species at the within-host level (e.g. (Choisy and de Roode 2010)), few have modeled these dynamics while considering both spatial scales (although see (Alizon and van Baalen 2008)).

Metacommunity models, in a sense, ask more general and larger-scale questions than epidemiological models, though they are similar in many ways (Box 3.1). This theory makes predictions regarding the roles of local processes and dispersal among localities in partitioning regional (γ) diversity into local habitats (α-diversity) and how this local diversity varies across space (e.g. β-diversity). Metacommunity models explore general trends, such as how dispersal changes the coexistence of competing species or how dispersal and patch heterogeneity affect β-diversity across the landscape. In other words, metacommunity theory is concerned with the *structure* of communities and how this structure varies across space, in response to various mechanisms. Below, I will highlight many specific practical and general applications that stem

from considering these types of questions with symbionts.

3.4.2 Practical applications

Linking metacommunity theory and the study of symbionts has many practical applications. In general, understanding the structure of symbiont metacommunities, as well as the underlying causes of this structure, could offer new insights into how symbionts affect host health. The diversity of symbionts within a host (α-diversity) might have direct effects on, or be correlated with, host health and general functioning. Understanding the turnover among symbiont communities (β-diversity) is also important, as the effect of symbiont communities on their hosts presumably changes as their compositions change (e.g. (Telfer et al. 2010)).

Incorporating ideas from the 'evolving metacommunity' concept is particularly relevant to the study of symbiont communities due to the overlap of ecological and evolutionary timescales in these systems (De Meester 2011). While the details of this framework will be discussed below, the theory could be used to understand or model, for example, the distribution of symbiont haplotypes among hosts in response to gradients in host genetic diversity, considering the role of host-symbiont genotype interactions in symbiont fitness. Also, this framework would be relevant for studying kin selection dynamics in the evolution of parasitic symbiont virulence. For example it is important to understand how multiple related strains of parasite interact within a host and how these intra-host interactions and inter-host transmission dynamics influence the evolution of parasite traits (e.g. (Alizon and Lion 2011)).

3.4.3 General ecological applications

Linking metacommunity theory and symbiont evolutionary ecology also has implications

for understanding the structure of ecological communities in general. For instance, different ecological and evolutionary processes might influence the structure of symbiont metacommunities compared to those of free-living organisms. As revealed by a large meta-analysis, most free-living organism metacommunity structures tend to show strong effects of species sorting and mass effects, but both local and regional components of assembly are necessary to explain the majority of variation (Cottenie 2005). Symbiont metacommunities could be compiled and compared to these data in order to determine if symbiont metacommunities are structured similarly to those of free-living organisms.

Research is already attempting to discern the primary drivers of human microbial symbiont community composition using metacommunity theory. For example, Lindström and Lengenheder (Lindström and Langenheder 2011) recently applied the predictions of mechanistic-based metacommunity theory to discuss general trends in bacterial community assembly across a wide range of systems, including the human body. Additionally, a model-fitting study used neutral metacommunity theory to explore how microbial α-diversity of the human lungs and digestive tract is structured (Sloan et al. 2006). Neutral metacommunity theory has also been used to speculate on the structuring mechanisms of parasitic helminth communities (Poulin 2004). However, these studies could be deepened and extended to encapsulate the full range of metacommunity theory applications (Box 3.1).

Different structuring mechanisms might act at different spatial scales for symbiont metacommunities compared to free-living organisms. Svensson-Coelho and Ricklefs (Svensson-Coelho and Ricklefs 2011) recently explored how avian host phylogeography relates to haemosporidian parasite community structure across the Lesser Antilles. Host genetic diversity between islands does not predict β-diversity of these parasites. Furthermore neither mosquito nor

bird β-diversity correlated with parasite β-diversity. Here, the spatial extent of the data analysis might overreach the scale at which the primary parasite community structuring processes function. Hierarchical analysis of the parasite metacommunity structure could reveal the relative roles of different structuring processes at different spatial extents (Presley and Willig 2010, Presley et al. 2010; Box 3.1).

Symbiont communities could be used to test metacommunity predictions by using pattern-based metacommunity approaches (Box 3.1). De Meester (2011) speculates that small organisms (e.g. environmental bacteria) might show an even stronger signal of species sorting due to their high dispersal capacity and rapid population growth, allowing for the simultaneous influence of dispersal and evolution to match organisms to their environmental optimums across space (see below). Symbiont dispersal and transmission rates, however, are linked with their hosts' dispersal rates to varying degrees (e.g. (Blasco-Costa et al. 2012)). Thus, symbiont metacommunities could be divided into 'meta-ensembles' (*sensu* (López-González et al. 2012b)) of high and low dispersing varieties and the respective structures could be used to test the assumed role of dispersal in the four predominant metacommunity paradigms.

3.4.4 General evolutionary applications

The 'evolving metacommunity' concept considers how gene flow and local adaptation can alter local species interactions and lead to either regional coexistence or monopolization (Urban et al. 2008). Consider a simple scenario in which competing species are locally adapted but can disperse to patches with different local conditions (Loeuille and Leibold 2008, Urban and De Meester 2009). Species A invades an empty patch and becomes locally adapted, preventing the establishment of its competitor, Species B. This is an example of local monopolization

facilitated by adaptation. However, if local conditions vary between patches, gene flow from Species A's new patch to its source could lead to maladaptation in the source. This could then allow Species B to invade Species A's source patch. The maladaptation of Species A thus facilitates regional coexistence with Species B. Similar theory can easily be applied to and tested with symbionts, especially considering their rapid population growth rates and short-term evolutionary dynamics (e.g. (De Meester 2011)). However, in host-symbiont systems, the relative rate of gene flow of hosts versus symbionts is important for symbiont adaptation, especially in antagonistic coevolution scenarios (Greischar and Koskella 2007, Hoeksema and Forde 2008). Thus, understanding the relation between host population gene flow and symbiont gene flow among individual hosts will be essential for predicting and interpreting evolutionary symbiont metacommunity patterns.

Recent conceptual models predict that intermediate levels of gene flow among patches in a metacommunity will maximize the adaptation of community and population traits to the local environment via species sorting and natural selection (Urban et al. 2008). In other words, regional biodiversity might be hindered by too little or too much dispersal, due to monopolization or maladaptation. Therefore, understanding the level of symbiont gene flow among hosts will be critical to predicting the conditions under which coexistence or monopolization will be favored. Future studies could explore how within-host symbiont community dynamics interact with variable dispersal or transmission rates to influence among-host community structural patterns (Box 3.2).

3.5 Conclusion

Symbiont communities are ripe for a merger with existing and emerging research on

metacommunity dynamics. This would simultaneously allow for testing theory and understanding the effects of symbionts on host health, whether positive or negative. Utilizing symbionts in metacommunity research and applying concepts garnered by metacommunity studies to symbiont eco-evolution could aid in our general understanding of spatial ecological dynamics, help determine the processes and spatial scales most relevant to the structuring of symbiont communities, guide studies related to symbiont trait evolution, and offer a framework for exploring the effects of symbiont communities on disease across space. Metacommunity ecology is a new and burgeoning field, and incorporating symbionts into its theoretical and experimental repertoire is bound to benefit ecologists, evolutionary biologists and medical scientists alike.

Box 3.1. Standard applications of metacommunity theory

Generally, ecological metacommunity theory has been approached in two ways (Figure 1). The discussion that follows is not comprehensive but, rather, is meant to introduce readers to the various ways metacommunity theory has been used to date.

Mechanism-based Approach

The mechanism-based approach develops and tests theoretical models that generate predictions about how the regional species pool partitions into local habitats, and how communities vary across space (Figure 1). Metacommunity models can be separated into four paradigms that mainly differ in the role of patch heterogeneity and the timing and effect of dispersal on local dynamics (Leibold et al. 2004b, Holyoak et al. 2005) (Figure 1). Most metacommunity models consider space implicitly, where species can differ in dispersal rates, but

it is assumed that all patches are colonized with equal probability. More recent models, however, explore the implications of spatially explicit patch distributions (e.g. (Pillai et al. 2011)). These paradigms have been tested using a variety of study systems and procedures (reviewed in (Logue et al. 2011)). Data generally show that community composition patterns are best explained by integrating multiple paradigms.

Mechanism-based theory has been used to investigate many real-word issues. Recent examples include: examining how habitat destruction alters food web complexity (Pillai et al. 2011); using patch connectedness to determine the best conservation methods (Economo 2011); and evaluating how climate change might alter community compositions (Urban et al. 2011).

Pattern-based Approach

The pattern-based approach examines the structure of natural metacommunities and evaluates the influence of particular environmental gradients in creating those structures (Figure 1). This approach relies on ordinations and null models to determine which 'idealized' metacommunity structures, if any, best fit to the observed data (Leibold and Mikkelson 2002a, Presley et al. 2010b). Then, canonical correspondence analysis is used to test what natural biotic and abiotic factors might lead to the observed structure (Ter Braak and Prentice 1988, López-González et al. 2012b).

As a cartoon example, metacommunities (a) and (b) in Figure 1 have Clementsian structure, in which discrete communities replace each other along a gradient, whereas metacommunity (c) has Gleasonian structure, in which each species responds individually to any gradient. There are three distinct sub-metacommunities, or compartments, within metacommunity (b), each of which has a nested structure. Interestingly, in natural

metacommunities, compartments are probably common, and each compartment can have a unique structure that responds to different biotic and/or abiotic gradients (Presley and Willig 2010b, López-González et al. 2012b). Thus, hierarchical analyses can determine which mechanisms are important at which spatial scales (Presley et al. 2010b).

Box 3.2. Symbiont metacommunity research topics at multiple spatial scales

Here, I propose ecological and evolutionary research topics that emerge from a consideration of symbiont metacommunities at various spatial scales (Figure 2). These questions are diverse, relevant to the general understanding of ecological communities as well as the effects of symbionts on host health. However, this list is surely not exhaustive.

Some important general questions:

- Which metacommunity processes are most important at the different spatial scales of symbiont metacommunities?
- How might metacommunity processes that occur at different spatial scales interact to affect symbiont community composition?
- Is local and/or regional symbiont community structure (e.g. richness, evenness) or metacommunity structure a reliable predictor of disease risk or overall host population health?

Intra-host:

- How does the level of symbiont gene flow between areas within the host (e.g. organs) affect the emergence of novel and/or pathogenic varieties?

- Is the rate and outcome of symbiont evolution influenced by the resident symbiont community composition within different host compartments?
- Could controlling symbiont dispersal to certain compartments and/or a spatially directed use of pro-microorganism treatments combat the intra-host evolution of pathogenic strains?

Inter-host (in a host population):

- How does heterogeneity within a host population (e.g. genetic diversity) influence symbiont gene flow among hosts?
- Does host heterogeneity influence symbiont community composition within and among hosts and, therefore, the potential for epidemics?

Inter-host sub-population (in a host metapopulation):

- How does host dispersal between sub-populations augment local symbiont community composition, and what consequences does this have for population-level disease risk?
- Considering the effect on host health, how does symbiont β-diversity across host sub-populations affect host metapopulation stability?

CHAPTER 4

USING MULTI-SPECES OCCUPANCY MODELS TO IMPROVE THE CHARACTERIZATION AND UNDERSTANDING OF METACOMMUNITY STRUCTURE

4.1 Abstract

Two of the most prominent frameworks to develop in ecology over the past decade are metacommunity ecology, which seeks to characterize multi-species distributions across space, and occupancy modeling, which corrects for imperfect detection in an effort to better understand species occurrence patterns Although their goals are complementary, metacommunity theory and statistical occupancy modeling methods have developed independently. For instance, the elements of metacommunity structure (EMS) framework uses species occurrence data to classify metacommunity structure and link it to underlying environmental gradients. While the efficacy of this approach relies on the quality of the data, few studies have considered how imperfect detection – which is widespread in ecological surveys and the major focus of occupancy modeling – affects the outcome. Here we introduce a framework that integrates multi-species occupancy models with the current EMS framework – detection error-corrected EMS (DECEMS). This method offers two distinct advantages. First, DECEMS reduces bias in characterizing metacommunity structure by using repeat surveys and occupancy models to disentangle species-specific occupancy and detection probabilities, ultimately bringing metacommunity structure classification into a more probabilistic framework. Second, occupancy modeling allows estimation of species-specific responses to environmental covariates, which will increase our ability to link species-level effects to metacommunity-wide patterns. After reviewing the EMS framework, we introduce a simple multi-species occupancy model and show how DECEMS can work in practice, highlighting that detection error often causes EMS to assign

incorrect structures. To emphasize the broader applicability of this approach, we further illustrate that DECEMS can reduce the rate of structure misclassification by more than 20% in some cases, even proving useful when detection error rates are quite low (~10%). Integrating occupancy models and the EMS framework will lead to more accurate descriptions of metacommunity structure and to a greater understanding of the mechanisms by which different structures arise.

4.2 Introduction

Over the last decade, metacommunity ecology has integrated ecological theory across spatial scales in an effort to better understand local and regional community dynamics. In particular, metacommunity theory explores how regional processes such as dispersal combine with local species interactions to affect species coexistence across scales (Leibold et al. 2004a, Holyoak et al. 2005, Chase 2005a). Metacommunity research often seeks to empirically characterize multi-species spatial distribution patterns with the ultimate goal of linking these patterns to underlying biotic and abiotic gradients or processes. In this effort, the elements of metacommunity structure (EMS) framework (Leibold and Mikkelson 2002b, Presley et al. 2010a) was developed as a set of analytical tools to identify and classify structural patterns in community data sets. The EMS framework has been used to better understand metacommunity dynamics across a variety of habitats and taxa (Presley and Willig 2010a, López-González et al. 2012a, Henriques-Silva et al. 2013, Meynard et al. 2013, Richgels et al. 2013, Erős et al. 2014). For instance, by associating unique community structures with areas of endemism and historical refugia, de la Sancha et al. (2014) showed that the structures of South American Atlantic Forest small-mammal communities were most consistent with trends in historical biogeography rather than current anthropogenic impacts. However, the value of the EMS framework could be

substantially enhanced by synthesizing its foundations of pattern-detection with recent advancements in the statistical modeling of species' occupancy.

The EMS framework uses observed species occurrence data aggregated across 'patches' of habitat in a metacommunity (e.g. field sampling sites), which are compiled into a site-by-species incidence matrix. Three summary statistics (coherence, turnover, and boundary clumping) are derived from this matrix to determine which of six core categorical structures the metacommunity exhibits (Presley et al. 2010a). Based on the specific structure observed, inferences can be made as to how the metacommunity assembles, and further statistical analyses help to associate the observed structure with a dominant environmental gradient (e.g. elevation, landcover, etc.). However, the efficacy of this pattern-to-process approach depends heavily upon the quality of the data (Gotelli and Graves 1996, Ulrich and Gotelli 2013); problems with species detection could lead to incomplete incidence matrices and inaccurate assessments of metacommunity structure. Fortunately, a decade of advancements in occupancy modeling has led to powerful methods to overcome problems such as species detectability (MacKenzie et al. 2002, Royle and Dorazio 2008, Dorazio et al. 2010, Burton et al. 2012).

Occupancy models rely on repeated sampling surveys to distinguish between the probability of a species occurring at a site and the probability of a species being detected at a site in which it occurs (MacKenzie et al. 2002, Royle and Dorazio 2008). This approach allows for an estimation of 'true' occupancy at each sampled site. Multi-species occupancy models also estimate species-specific environmental covariate effects on occurrence probability (Dorazio et al. 2010, Burton et al. 2012). This means that community-level distribution patterns could be associated with species-specific responses to environmental gradients, an important advancement that is lacking in the EMS framework. Thus, by providing better estimates of occurrence and by

allowing estimation of species-specific covariate effects, occupancy modeling should improve our characterization and understanding of metacommunity structure.

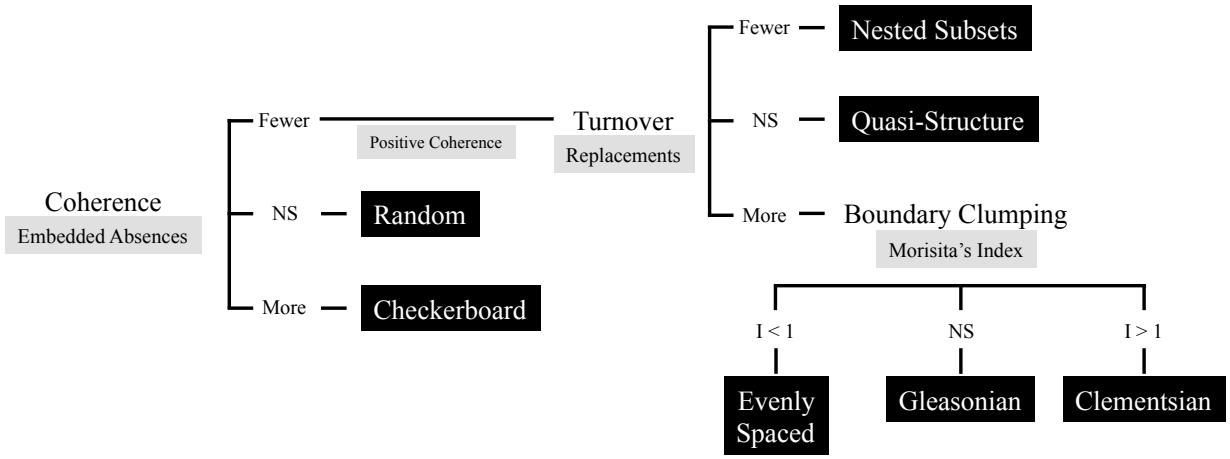
Here, we review current EMS methods and introduce a metacommunity framework that integrates multi-species occupancy models, which we term detection error-corrected EMS (DECEMS). This method aims to allow for more accurate categorization of structure and the estimation of species-specific responses to environmental gradients. To help illustrate the efficacy of this approach, we introduce a simulated case study, which shows how detection error can bias standard EMS approaches and how DECEMS can effectively overcome this problem. We also quantitatively demonstrate that DECEMS can significantly reduce the effect of imperfect detection on structure misclassification across a wide range of simulated detection error rates. By estimating species-specific covariate effects, DECEMS could also lead to more mechanistic understandings of metacommunity structure. We expect that this approach will facilitate a more complete understanding of metacommunities by improving metacommunity structure inference and by revealing how species-specific responses to environmental gradients might scale up to community-level patterns.

4.3 Elements of metacommunity structure

The EMS paradigm follows a step-wise procedure to determine which of six core categorical metacommunity structures are exhibited by a data set of species occurrences observed across multiple sites. Although the procedure can determine that no orderly structure exists (i.e. random structure), most empirical metacommunities tested thus far exhibit detectable structure. Here we provide an overview of these methods and discuss some problems that can arise.

Species occurrence data (or abundance data) are assembled into a site-by-species matrix and this matrix is ordinated, typically using reciprocal averaging. Reciprocal averaging is a type of correspondence analysis that uses an algorithm to generate ordination scores based on the sites' similarities in species composition and the species' similarities in distribution among sites. The original matrix is rearranged (i.e. ordinated) based on the primary ordination axis scores to group similar sites and similarly distributed species. This ordinated matrix theoretically represents how species assemblages are structured along a dominant environmental axis (i.e. gradient). For example, Mexican bat species form discrete assemblages that turnover along a humidity gradient (López-González et al. 2012a). Then, from the ordinated matrix, statistics are calculated to summarize the three elements of metacommunity structure: coherence, turnover, and boundary clumping. These statistics are used to assign one of six core categorical metacommunity structures (Figure 4.1, Leibold and Mikkelson 2002, Presley et al. 2010).

Figure 4.1. Flow chart used to determine metacommunity structure. Text in gray boxes represents the metric used to estimate the corresponding element of structure (e.g. embedded absences are used to determine coherence). NS = non-significant based on simulated null matrices.



The first metric of metacommunity structure is **coherence**, which reflects whether the majority of species in the metacommunity respond to the same axis of variation, often assumed to be an environmental gradient. Coherence is the foundation of structure as without either significant positive or negative coherence, the community is said to be randomly structured (i.e. species do not structure along a common axis of variation; Figure 4.1). Coherence is estimated using the number of embedded absences, which occur in areas of the matrix where a species is absent at a site in which it would be expected to occur based on the ordination. The observed number of embedded absences is then compared to a null distribution of embedded absences generated from ~1000 simulated matrices.

Significantly negative coherence is indicative of a checkerboard pattern, suggesting negative pair-wise species associations. However, if a metacommunity exhibits positive coherence, two more metrics are calculated to further describe the metacommunity structure: **turnover** and **boundary clumping**. Turnover represents how species composition changes along the theoretical environmental gradient, estimated using the number of species replacements observed in the ordinated matrix. Negative turnover (significantly fewer replacements than the null) is indicative of nested subsets (i.e. a core local assemblage with species subtractions along the gradient), while positive turnover represents more substantial shifting composition (i.e. species additions and subtractions to local communities along the gradient). Boundary clumping, estimated with Morisita's index, helps to further distinguish structures by determining whether distinct clusters of species aggregate along the gradient, or whether there is a more gradual, random shift in structure. For instance, with positive turnover, significant clumping would indicate Clementsian structure, where discrete species groups turnover along the gradient, whereas a lack of clumping would be consistent with Gleasonian structure, where species

respond idiosyncratically to the hypothetical gradient. If there is no significant turnover, various quasi-structures are assigned to the metacommunity depending on the trend observed in turnover and boundary clumping (*sensu* Presley et al. 2010).

After the metacommunity structure is characterized, researchers typically seek to determine which environmental covariate explains the primary axis of variation in the ordinated community. This analysis takes various forms. In most cases, the ordination score of the primary axis is extracted for each sampled site. Then univariate correlations are run for each covariate of interest against the ordination scores to explore how covariates might be responsible for structuring the metacommunity (e.g. Henriques-Silva et al. 2013, Meynard et al. 2013), although multivariate models could (and perhaps should) be used here as well. A complementary approach involves using canonical correspondence analysis (CCA) to relate the incidence matrix to multiple covariates simultaneously (e.g. López-González et al. 2012). Some recent studies combine one of these previously discussed analyses with a variance partitioning analysis of the incidence matrix to evaluate the relative contribution of classes of covariates, such as 'local' and 'spatial' or 'abiotic' and 'biotic' (Henriques-Silva et al. 2013, Dallas and Presley 2014). Finally, emerging research demonstrates the utility of combining hierarchical cluster analysis with CCA to determine which combination of sites represent distinct metacommunity compartments, and how environmental covariates might influence the formation of these compartments along a gradient (de la Sancha et al. 2014).

4.4 Challenges inherent to the EMS approach

4.4.1 A problem with detection error

The EMS approach relies on occurrence data, which often suffer from imperfect

detection owing to issues of sampling design and effort, low species' abundances, and idiosyncrasies in species' ecologies (e.g. cryptic or crepuscular organisms) (MacKenzie et al. 2002). Detection error can in turn influence the ordination of the community incidence matrix, the calculated EMS metrics, and the accuracy of structural inference based on null matrices. If species detection is imperfect, for example, the calculated number of embedded absences from the ordinated matrix may be overestimated, which could lead to misclassification errors (analogous to type II errors), where metacommunity structures are assigned incorrectly. Additionally, imperfect detection influences the form of null matrices, as most methods of null matrix generation utilize the raw data on row and/or column sums (Gotelli 2000, Ulrich and Gotelli 2013).

4.4.2 Inferring structuring mechanisms from covariates

Leibold and Mikkelson (2002) emphasized that the EMS methods can identify patterns, but cannot necessarily elucidate the processes that lead to pattern. For example, metacommunities that exhibit Gleasonian and Clementsian structure are hypothesized to be structured differently based on species-specific responses to a dominant environmental gradient (Clements 1916, Gleason 1926, Gilpin and Diamond 1982). Gleasonian structure is believed to arise from idiosyncratic species responses, whereas Clementsian structure arises from groups of species that respond similarly to each other but differently from other groups of species in the metacommunity. Alternatively, however, Clementsian structure could arise from negative associations between species pairs or groups that arise along the gradient (Gilpin and Diamond 1982). These mechanistic interpretations of structure remain speculative in the EMS paradigm given that these methods do not estimate species-specific covariate effects. Rather, EMS

methods tend to rely on correlating the ordination scores from a single axis of variation – which could be leaving out valuable information about community structure – to environmental covariates.

4.5 Linking EMS and multi-species occupancy models

The occupancy modeling framework can help ameliorate the issues addressed above by: (1) disentangling occupancy and detection probabilities of each species, and (2) estimating species-specific covariate effects, which can complement the use of ordination scores and will allow for empirical tests of hypotheses related to structuring mechanisms.

Occupancy models were developed to estimate a species' probability of occurring at a site while correcting for the fact that species may go undetected in a survey (MacKenzie et al. 2002, Royle and Dorazio 2008). These models use data from repeat surveys conducted in a time period during which the true occupancy status of a site is assumed to be constant (i.e. the occurring species are not transient). This allows one to disentangle detection and occurrence probabilities in order to estimate true occupancy at each site and obtain unbiased estimates of a species' response to environmental covariates. More recently, these models have been extended to multi-species and multi-time point (i.e. longitudinal) surveys (Dorazio et al. 2010, Burton et al. 2012). These models incorporate species-, site- and time-specific estimates of detection probability, occurrence probability, and covariate effects. Additionally, with longitudinal surveys, one can estimate probabilities of persistence at a site and colonization of previously unoccupied sites.

4.6 Formulating the multi-species occupancy model

We use a multi-species occupancy model with multiple surveys at each site over a single

time period (e.g. three surveys over a month). More complex models can be designed, but we wish to keep our analyses tractable to demonstrate the utility of occupancy models in the EMS paradigm with a simple example.

Let $z_{i,k}$ represent the true occurrence of species *i* at site *k*, where $z_{1,2} = 1$ means that Species 1 is present at Site 2. These $z_{i,k}$ values can be compiled into a 'true' metacommunity occurrence matrix, **Z**. True occurrence states arise as Bernoulli random variables with probability, $\psi_{i,k}$, the probability of occurrence:

$$z_{i,k} \sim Bern(\psi_{i,k})$$

We assume that the probability of occurrence, $\psi_{i,k}$, is related to a single continuous environmental covariate, though any number of covariates could be used in practice:

$$logit(\psi_{i,k}) = \beta_{i0} + \beta_i x_k,$$

where β_{i0} is the species-specific intercept, β_i is the effect of covariate *x* on species *i*, and x_k is the value of covariate *x* at site *k*.

We assume that multiple surveys are conducted at each site and observations are compiled into a species-by-site occurrence matrix, **Y**. For example, if Species 1 is observed in two out of three surveys at Site 2, $y_{1,2} = 2$. The number of observed occurrences of each species out of the total number of re-surveys at each site facilitates estimation of species-specific detection probabilities. Let p_i represent the probability of detection of species *i*. Although detection probability can be related to covariates similarly to occurrence probability, for simplicity we did not impose any such covariate effects.

The observed occurrences are thus binomially distributed, influenced by both the detection and occurrence probabilities and the number of surveys conducted at each site, J_k:

$$y_{i,k} \sim Binom(J_k, z_{i,k} p_i)$$

Here we adopt a Bayesian approach for inference and parameter estimation, utilizing Markov chain Monte Carlo (MCMC) sampling in which we iteratively sample from the posterior distribution of each true occupancy state, $z_{i,k}$. For each draw from the posterior, we then obtain a site-by-species incidence matrix $\mathbf{Z}_{post}$, which consists of the elements $z_{i,k}$.

For all of our simulations below, we used the open-source statistical software, *R* (R Core Team 2014). For Bayesian analyses we used the open-source software JAGS (Just Another Gibbs Sampler; http://mcmc-jags.sourceforge.net/). For metacommunity analyses, we used the *R* package *metacom* (Dallas 2014), which relies heavily on the *R* package *vegan* (Oksanen et al. 2013).

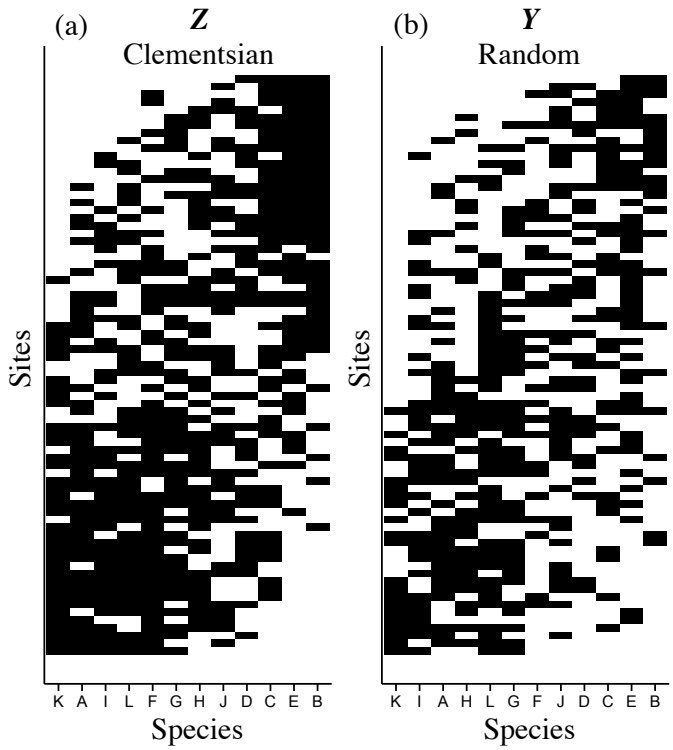
4.7 Integrating EMS and occupancy models: A simulated example of DECEMS

In this section we use a simulated case study to introduce detection error-corrected EMS (DECEMS), which integrates occupancy modeling and standard EMS. First, we simulated a metacommunity of 12 species and 75 sampled sites, assuming each site was sampled 3 times. In order to create a coherent metacommunity, we imposed a dominant environmental covariate to which species had variable responses. In this way, the community composition shifts along a gradient of the covariate, analogous to say, an elevational gradient. We set $\beta_{i0} = logit(0.60)$ for all species, and we assumed that species-specific covariate effects, β_i, follwed a normal distribution with mean=0 and standard deviation=1. Site-specific covariate values followed a normal distribution with mean=0 and standard deviation=2. To emphasize how detection errors can obscure true metacommunity structure, we assumed the species-specific probabilities of detection, $logit(p_i)$, followed a normal distribution with mean=$logit(0.5)$ and a standard deviation=0.75. This represents a community in which many species are difficult to detect, a

scenario under which an occupancy model might be most useful.

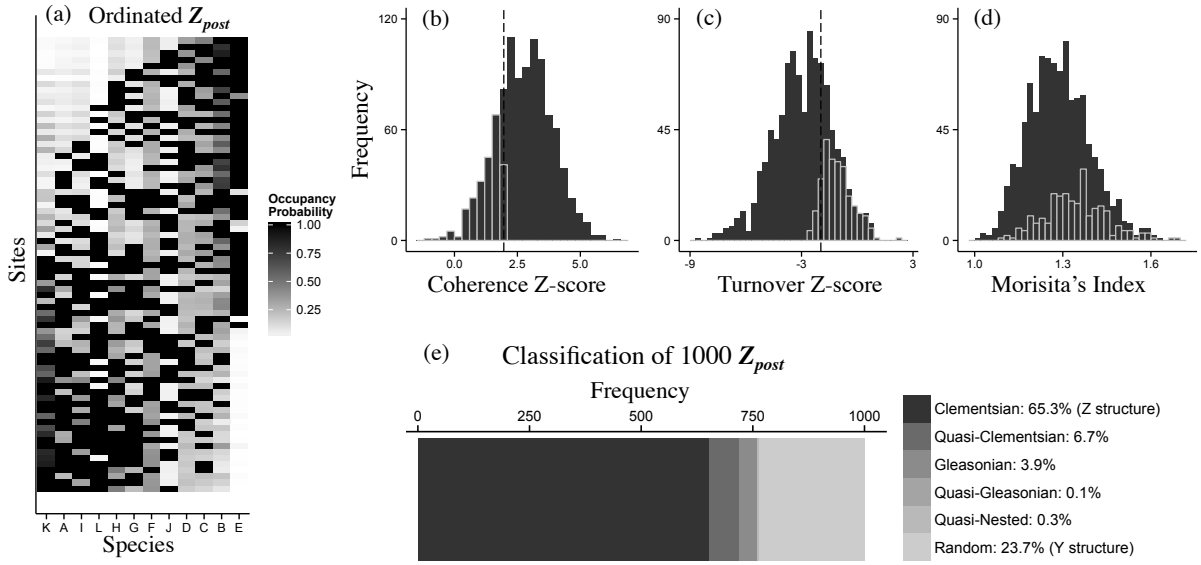
For this simulated metacommunity, we thus had a known occurrence matrix, **Z**, and an observed occurrence matrix, **Y**, and we used EMS methods to characterize metacommunity structure of **Z** compared to **Y**. In this example, the known metacommunity **Z** displayed Clementsian structure, while the observed metacommunity, **Y**, displayed random (i.e. no discernible) structure based on standard EMS methods (Figure 4.2). Here, detection error resulted in more embedded absences and a concomitant rearrangement of species and sites in the ordinated matrix. It is worth noting that following standard EMS methods, in this example, the only result is that the metacommunity displays random structure, which we know is incorrect. Next, we applied the occupancy model to estimate the true occupancy states **Z**, based on the observed data, **Y**. We used uninformative priors, ran the model with a 1000 iteration adaptive phase, followed by a 5000 iteration burn-in period. After the burn-in period, we ran the model for 10,000 iterations, thinning the MCMC chains by 10 iterations, for a final sample of 1000 $\mathbf{Z}_{post}$ across 3 MCMC chains. We assessed convergence using the potential scale reduction factor, $\hat{R}$ (Gelman 1996).

Figure 4.2. The ordinated form of the (a) known occurrence matrix **Z**, and the (b) corresponding observed occurrence matrix **Y** with imposed detection error from our simulated example.



For 1000 $\mathbf{Z}_{post}$ from the Bayesian occupancy model, we calculated metacommunity metrics, assembling psuedo-posterior distributions for each metric (Figure 4.3b-d). These are 'psuedo'-posteriors because we used frequentist statistics, based on standard null (i.e. randomized) matrix simulations, to calculate the significance of the metacommunity metrics for each $\mathbf{Z}_{post}$; however, these $\mathbf{Z}_{post}$ were estimated in the Bayesian framework. We thus determined the metacommunity structure of each $\mathbf{Z}_{post}$.

Figure 4.3. Data on metacommunity structure derived from occupancy model. (a) A heat map ordination of the $\mathbf{Z}_{post}$ matrices, based on the average site- and species-specific probability of occurrence across the 1000 iterations. These values were then used to ordinate the species and sites with detrended correspondence analysis. (b-d) Histograms of the metacommunity metrics for each $\mathbf{Z}_{post}$. Bars highlighted in light gray represent $\mathbf{Z}_{post}$ that show non-signficant coherence (i.e. random structure). Vertical dashed lines in (b-c) delineate the significance cut-off for the metrics' normalized z-scores at $\alpha = 0.05$. (e) Distribution of 1000 $\mathbf{Z}_{post}$ metacommunity structures estimated by the occupancy model.

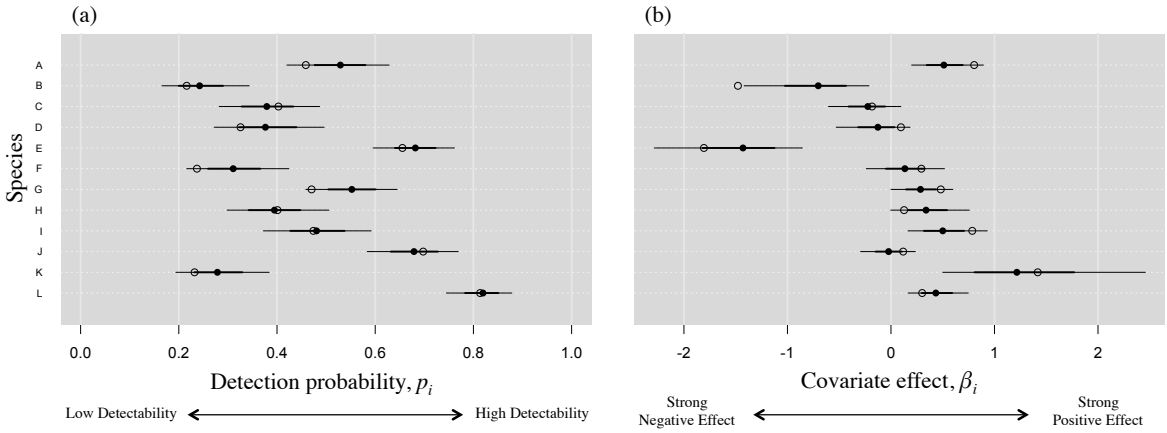


In this example, using DECEM we found that most (65.3%) $\mathbf{Z}_{post}$ matched the true metacommunity structure of $\mathbf{Z}$ (Clementsian), rather than the incorrect random metacommunity structure (23.7%) associated with observed $\mathbf{Y}$ (Figure 4.3e). An additional 6.7% of $\mathbf{Z}_{post}$ showed Quasi-Clementsian structure. We also created a new ordinated matrix based on the species- and site-specific probabilities of occurence, estimated as the proportion of occurrences observed across the 1000 $\mathbf{Z}_{post}$ (Figure 4.3a). Importantly, although the occupancy model does not find every $\mathbf{Z}_{post}$ to match $\mathbf{Z}$, this integration of methods now puts metacommunity structure into a probabilistic framework.

The model was also able to accurately estimate the species-specific probabilities of detection, p_i, and species-specific covariate effects, β_i (Figure 4.4). Having these species-

specific effect estimates allows us to see how species-level responses can scale up to metacommunity-wide patterns. For instance, the Clementsian structure observed in **Z** is likely driven by a few species that have strong responses (either positive or negative) to the dominant covariate. This would preclude them from habitats with more extreme covariate values, leading to a clumped distribution. The occupancy model shows that, in our example, Species B, E, and K have relatively strong responses compared to the rest of the metacommunity members. This matches the pattern observed in the metacommunity ordination, showing that indeed these three species' responses to the underlying gradient tend to drive the Clementsian pattern.

Figure 4.4. Values of species-specific probabilities of detection, p_i, and species-specific covariate effects, β_i, as estimated by the occupancy model. Filled circles mark the medians of the posterior distributions of each parameter, while open circles mark the true (simulated) values. Thicker and thinner lines represent the 68% and 95% credible intervals of the estimates, respectively.



This example shows that detection error can bias the assignment of metacommunity structure and that, at least in this case, the occupancy model can provide a more accurate picture of metacommunity structure. Furthermore, it demonstrates the utility of estimating species-specific covariate effects, which helps us elucidate how species-level responses can influence

overall metacommunity structure.

4.8 Effect of detection error correction on EMS misclassification rate

Although an occupancy model was useful in the simulated example above, we wanted to quantitatively explore whether occupancy modeling reduces the misclassification rate in assigning metacommunity structure across different p_i distributions. In other words, we wanted to answer the question of whether occupancy models are useful generally or only when detection errors are more extreme.

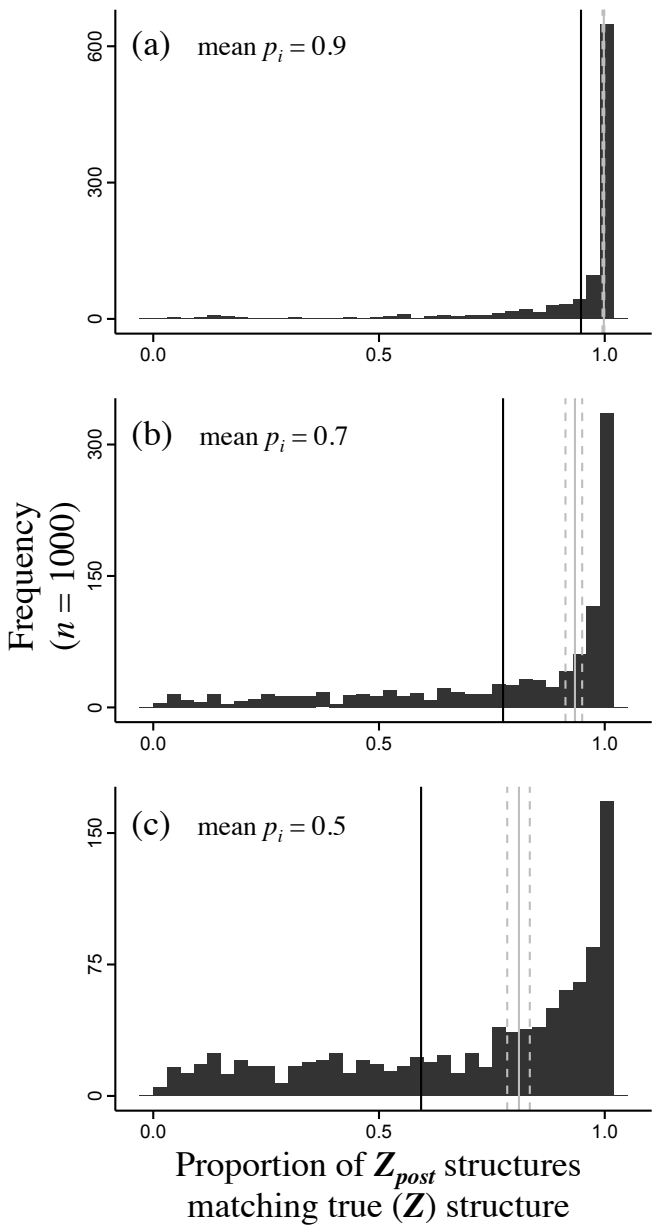
To address this question, we simulated 1000 unique metacommunities again using a global pool of 12 species and 75 sites surveyed 3 times each. We assumed a dominant environmental covariate to which species responded, in order to achieve coherence in most cases. In order to simulate different metacommunity structures, we randomly varied the distribution of species-specific covariate effects and covariate values (e.g. by varying the distribution type – normal or uniform – and variability – standard deviation or range).
For each simulated metacommunity ($n = 1000$) we thus had a known occurrence matrix, $\mathbf{Z}$, and an observed occurrence matrix, $\mathbf{Y}$. We used EMS methods to categorize the metacommunity structure for each $\mathbf{Z}$ and $\mathbf{Y}$. For each metacommunity, we then used an occupancy model to estimate the posterior distribution of $\mathbf{Z}$ by drawing 500 iterations of the posterior estimated occurrences, $\mathbf{Z}_{post}$, and determining the metacommunity structure for each $\mathbf{Z}_{post}$. We conducted this full simulation three times, fixing the mean probability of detection at three values: 0.9, 0.7, or 0.5, with a standard deviation of 0.75 (e.g. $logit(p_i) \sim N(logit(0.9),0.75)$.

To determine whether the occupancy model reduced bias in estimating metacommunity structure, we compared the percentage of cases in which the assigned metacommunity structure of observed occurrence matrix, $\mathbf{Y}$, deviated from known occurrence matrix, $\mathbf{Z}$ – a point estimate

– to the median (and its bootstrapped 95% confidence interval) of the distribution of percentages of times that the structure of $\mathbf{Z}_{post}$ matched that of $\mathbf{Z}$. Effectively, this tested whether correcting for imperfect detection significantly reduced the misclassification rate of the 1000 simulated metacommunities compared to the standard EMS approach.

For all three values of mean p_i, the occupancy model significantly reduced bias in characterizing metacommunity structure (Figure 5). Even when the average species-specific detection probability was very high (i.e. mean $p_i = 0.9$), detection error resulted in 94.7% of observed occurrence matrices' ($\mathbf{Y}$) structures matching the true occurrence matrices' ($\mathbf{Z}$) structures, meaning a 5.3% misclassification rate. However, using the occupancy model, the median (and 95% CI) of the proportion of posterior occurrence matrix estimates $\mathbf{Z}_{post}$ whose structure matched the true $\mathbf{Z}$ structure was 99.8% (99.5% - 99.9%), showing a significant improvement and reducing the misclassification rate to < 1% (Figure 4.5a). Furthermore, the benefit of using the occupancy model increased as the mean probability of detection in the community decreased (Figure 4.5b-c). Specifically, for mean $p_i = 0.7$, the structure of $\mathbf{Y}$ matched that of $\mathbf{Z}$ 77.5% of the time, while the median for $\mathbf{Z}_{post}$ matching $\mathbf{Z}$ was 93.4% (91.3% - 95.0%); for mean $p_i = 0.5$, the structure of $\mathbf{Y}$ matched that of $\mathbf{Z}$ 59.3% of the time, while the median for $\mathbf{Z}_{post}$ matching $\mathbf{Z}$ was 81.0% (78.4% - 83.4%).

Figure 4.5. Determining if an occupancy model reduces bias in assigning metacommunity structure. Black vertical lines represent the point estimate of the proportion of 1000 simulations for which the observed metacommunity, **Y**, structure matched the known metacommunity, **Z**, structure. Solid and dashed gray vertical lines represent the median and bootstrapped 95% confidence intervals for the proportion of 500 estimated $\mathbf{Z}_{post}$ structures that match the known metacommunity, **Z**, structure. Parameters: (a) mean $p_i = 0.9$, (b) mean $p_i = 0.7$, (c) mean $p_i = 0.5$.



These results indicate that, somewhat surprisingly, incorporating an occupancy model increases the accuracy in assigning metacommunity structure, even when there are relatively high detection probabilities (~90%) among species. Using this approach could be especially important in cases of low average detection, for instance in microbial or symbiont communities (e.g. Mihaljevic 2012), in which species might be cryptic due to small size or aggregation patterns among hosts (e.g. negative binomial distribution).

4.9 Discussion

Integrating multi-species occupancy models into the EMS framework can effectively reduce bias in assigning metacommunity structure when there is error in species detection, which is a ubiquitous problem in occurrence data (MacKenzie et al. 2002). We found that integrating occupancy modeling with EMS (i.e. DECEMS) can lead to striking reductions in the rates of metacommunity structure misclassification that results from imperfect detection, even when detection error was quite low. This is an important improvement that should ensure the best possible classification of metacommunity structure from community data sets. Given that an occupancy model estimates species-specific covariate effects, this method can also be used to better inform how species-level responses can scale up to affect metacommunity-wide patterns of occurrence. This helps to address a key gap within EMS by improving our ability to link metacommunity patterns to species-level processes.

Based on these findings, we suggest that metacommunity ecologists will often benefit from using occupancy models in their assessment of metacommunity structure and in determining the environmental covariates that might lead to structure. This method should prove particularly useful in cases where species detection is known to be problematic, although our

results show that DECEMS improves EMS performance even for data sets with little detection error. Occupancy modeling requires repeat surveys over a time period in which it can be safely assumed that community member composition is not changing. This design allows for the estimation of species-specific detection probabilities. The design of repeat surveys will depend on the biology of the system and the available resources. For instance, a researcher could conduct two surveys per week (if appropriate), or multiple observers could conduct independent surveys on the same day. Thus collecting data to accommodate an occupancy model does not necessarily have to increase effort, and the end result is improved accuracy.

Future extensions to the occupancy-modeling framework presented here could further improve how we assess the influence of environmental covariates on metacommunity structure. For instance, Jackson et al. (2012) presented a maximum likelihood method using multi-level models – for which occupancy models are a specific example – to determine how environmental variation leads to changes in community composition. These authors showed that estimating species-specific covariate effects can outperform common methods used to assess the influence of environmental covariates, such as canonical correspondence analysis (CCA) and nonmetric multidimensional scaling (NMDS). The methods presented by Jackson et al. (2012) could easily be integrated into the occupancy-modeling framework with the additional benefit of simultaneously estimating (and therefore correcting for) species-specific detection probabilities. In this way, the same model could simultaneously estimate metacommunity structure and determine the relative influences of environmental drivers.

Occupancy modeling could also be used in theoretical metacommunity studies to help us understand when and where metacommunity structures might arise. Given that an occupancy model is able to estimate species-specific covariate effects, one could explore hypotheses about

how metacommunity structures arise in a quantitative framework. For instance simulation could be used to ask how the distribution of species-specific covariate effects and the distribution of covariate values observed among sampled sites affect resulting metacommunity structures. Such studies would strengthen the linkages between pattern-based metacommunity studies and mechanism-based theory.

The benefits of correcting for imperfect detection are increasingly appreciated in the fields of ecology and biogeography (Royle et al. 2012, Fitzpatrick et al. 2013, Iknayan et al. 2013, Lahoz-Monfort et al. 2014). Here we have shown that integrating occupancy modeling and EMS (i.e. DECEMS) can improve the accuracy of metacommunity structure classification, an important step towards understanding where and why certain structures emerge. We propose that a continued merger of the fields of occupancy modeling and metacommunity ecology should enrich and deepen our study of populations and communities across spatial scales.

CHAPTER 5

SYMBIONT METACOMMUNITIES THROUGH TIME: THE INFLUENCE OF ALTERNATIVE HOST COMPOSITION AND ABIOTIC HABITAT CHARACTERISTICS ON THE SYMBIONT COMMUNITIES OF METAMORPHIC FROGS

5.1 Abstract

Understanding the drivers of symbiont community composition across space and time has implications not only for testing and building ecological theory, but also for predicting how symbiont communities affect host health and overall symbiont transmission. Here we analyze a large-scale data set of amphibian symbiont communities surveyed in California wetlands over a four-year period. We use newly developed and novel metacommunity tools to characterize how symbiont metacommunity structure changes over time. Our statistical methods also allow us to identify the environmental drivers of symbiont community composition across space by estimating species-specific responses to biotic and abiotic covariates. We found that symbiont metacommunity structure is not consistent over time, showing random, nested, and Clementsian structures. The dominant drivers of symbiont community composition also changed from year to year but typically consisted of characteristics of the symbionts' intermediate host communities and geographic features of the wetlands. Furthermore, within years, symbiont responses to environmental gradients were typically not coherent, meaning that many species were affected by unique covariates rather than responding to a similar axis of variation. Our results highlight that although a symbiont community may aggregate in the same host, species in the community can be affected by unique aspects of the environment that influence colonization and subsequent transmission among hosts. Thus, predicting symbiont composition in this system, and likely in

many others, will require integrating across species-specific responses to various environmental gradients, emphasizing the utility of our newly developed methods.

5.2 Introduction

Evidence from various human and wildlife host-symbiont systems shows that shifts in symbiont community composition at various spatial scales can result in concomitant changes to host health and symbiont transmission patterns. For instance, re-establishment of a more diverse gut microflora via fecal transplantation can cure recurrent *Clostridium difficile* infections and relieve a variety of other ailments in humans (van Nood et al. 2013, Youngster et al. 2014). At larger spatial scales, symbiont composition can alter patterns of transmission. For example, the transmission of a highly virulent trematode of amphibians is significantly reduced in wetlands with more speciose parasite communities, most likely due to stronger intra-host competitive dynamics compared to those of wetlands with depauperate parasite communities (Johnson et al. 2013a). Given these types of patterns, understanding the drivers of symbiont community composition is an important goal, one to which the field of community ecology has much to offer (Pedersen and Fenton 2006, Mihaljevic 2012).

Symbiont communities are unique but well suited model systems for exploring general ecological patterns, with the applied benefit of informing host health and symbiont transmission dynamics (Dove 2006, Poulin 2007, Johnson and Hoverman 2012b, Fierer et al. 2012). For instance, understanding how symbiont community composition changes across space in response to both host population/community composition and exogenous environmental drivers is relevant to both general ecological theory and applied disease ecology (e.g. Morand and Poulin 1998, Krasnov and Vinarski 2008, Krasnov et al. 2010, Poulin et al. 2011). Fortunately, the field of

metacommunity ecology offers theory and analytical tools that are particularly relevant for disentangling the relative influence of various factors affecting species composition across space (Leibold et al. 2004b, Holyoak et al. 2005, Chase 2005b).

Symbiont communities can be considered metacommunities at various spatial scales, allowing for the application of metacommunity tools to answer a variety of questions (Mihaljevic 2012). For instance, Richgels et al. (2013) used the elements of metacommunity structure (EMS) framework (Leibold and Mikkelson 2002b, Presley et al. 2010b) to show how snail trematode parasite communities shift across California wetlands in response to local (i.e. wetland-level) and regional (i.e. among-wetland) habitat characteristics, suggesting that these communities are strongly filtered by local environmental attributes. At a larger spatial scale, Dallas and Presley (2014) used EMS methods and variance partitioning to determine how host species traits and phylogenetic relationships influence the structuring of parasite communities within desert rodents. These recent studies highlight the utility of both metacommunity concepts and statistical approaches for exploring how symbiont communities are structured across space.

In this study, we analyze a large-scale, multi-year data set of Pacific chorus frog (*Pseudacris regilla*) symbionts surveyed across wetlands in the San Francisco Bay Area of California. With this unique data set, we ask whether symbiont metacommunity structure changes from year to year and whether the factors that influence composition are consistent among years. To address these questions, we implement a recently developed method to characterize metacommunity structures, detection error-corrected elements of metacommunity structure (DECEMS) (Mihaljevic et al. *in review*). In combination with DECEMS, we use a novel statistical routine to investigate the factors driving symbiont compositional change across space. By using metacommunity theory and tools, we are able to address basic, yet highly

relevant, ecological questions about how symbiont communities structure across space and time. Furthermore, the methods we employ should prove informative to the study of metacommunities across a variety of taxa and habitats.

5.3 Methods

5.3.1 Study System

We surveyed wetlands across the Bay Area of California during the summers (July-August) of 2009 -2012 following the sampling design presented in Richgels et al. (2013) and Johnson et al. (2013b). These wetlands can harbor up to six amphibian species, five gastropod species, and many symbiont species that utilize multiple intermediate and definite host species within the wetlands. Here, we focus on the symbiont communities of the Pacific chorus frog, *P. regilla*.

During the aquatic larval life stage, these amphibians can acquire a variety of trematode and protozoan symbionts (Table 5.1), whose affects range from benign to severely pathogenic. The trematode species have complex life cycles that require multiple host species to complete reproductive and non-reproductive life stages. A typical life cycle of the trematodes found in *P. regilla* consists of a reproductive stage within a definitive host, typically a mammal or bird, which deposits trematode eggs into the wetland via feces. The eggs hatch and the next life stage infects a gastropod host, in which the trematodes reproduce asexually. A free-swimming life stage emerges from the gastropod host and encounters and infects another gastropod, insect, or amphibian host within the aquatic environment. These amphibians later metamorphose and are eaten by the trematode's definitive host to complete the life cycle. Typically, these trematodes

are more host-specific in the gastropod intermediate host, but are more generalist at later life stages. The trematode species tend to associate with different organs within the amphibian host.

Table 5.1. Classification of symbionts found in the field study.

Abbreviation	Kingdom	Phylum	Class	Order	Family	Genus	Species	Lifestage
Alar	Animalia	Platyhelminthes	Trematoda	Strigeidida	Diplostomatidae	Alaria	*Alaria spp.*	Trematode mesocercaria
Fib	Animalia	Platyhelminthes	Trematoda	Strigeidida	Diplostomatidae	Fibricola	*Fibricola spp.*	Trematode metacercaria
Glob	Animalia	Platyhelminthes	Trematoda	Plagiorchiida	Cephalogonimidae	Cephalogonimus	*Cephalogonimus spp.*	Trematode metacercaria
Echi	Animalia	Platyhelminthes	Trematoda	Echinostomida	Echinostomatidae	Echinostoma	*Echinostoma spp.*	Trematode metacercaria
Mano	Animalia	Platyhelminthes	Trematoda	Plagiorchiida	Ochetosomatidae	Manodistomum	*Manodistomum syntomentera*	Trematode metacercaria
Nyct	Protozoa	Ciliophora	Heterotrichea	Heterotrichida	Nyctotheridae	Nyctotherus	*Nyctotherus spp.*	Trophozoite
Opal	Chromalveolata	Heterokontophyta	Opalinea	Opalinida	Opalinidae	Opalina	*Opalina spp.*	Trophont
Rib	Animalia	Platyhelminthes	Trematoda	Echinostomida	Psilostomatidae	Ribeiroia	*Ribeiroia ondatrae*	Trematode metacercaria

The two protists that are associated with *P. regilla* are both located in the intestinal tract. Both species are directly transmitted, whereby adult frogs defecate eggs into the wetland that larval amphibians ingest.

5.3.2 Field Surveys

During field surveys of each wetland, recently metamorphosed *P. regilla* were collected from the wetland perimeter and sent to the University of Colorado at Boulder for necropsy examination and identification of symbiont species (Johnson et al. 2013b). Additionally, many site-level characteristics were measured and included as covariates of symbiont occupancy (Table 5.2). While the details of these methods are laid out in Richgels et al. (2013), in brief, a variety of geographical features, including latitude and longitude, elevation, slope, and aspect were recorded or later derived from geographic information systems. Local abiotic wetland characteristics, such as wetland area, water conductivity, dissolved oxygen, and total dissolved solids were also measured. We also surveyed local biotic attributes, such as the percentage of the shoreline vegetated, as well as the local community assemblages of amphibians and snails.

We derived four separate covariates from the occupancy and abundances of amphibian and snail species at the wetlands. For each year, and then again for all years' data combined, we ordinated the occupancy (pres/abs) of the amphibian species and, separately, of the snail species

community data using redundancy analysis (RDA). For each wetland we extracted the primary and secondary axes scores: Amph_RA1, Amph_RA2, Snails_RA1, and Snails_RA2. Similarly, we used abundance data in a non-metric multi-dimensional scaling (NMDS) analysis, and extracted the primary and secondary axes scores for each site: Amph_MDS1, Amph_MDS2, Snails_MDS1, and Snails_MDS2.

Table 5.2. Covariates used in our study and the transformations of the data.

Covariate	**Description**	**Transformation**
Lat	Latitude of the wetland	-
Long	Longitude of the wetland	-
Elev	Elevation of the wetland	Square root
Slope	Slope, calculated from USGS digital elevation data set (30m x 30m)	Square root
Aspect	Aspect, calculated from USGS digital elevation data set (30m x 30m)	-
Hydro	Wetland hydroperiod (nominal values: Permanent, Semi-permanent, Ephemeral)	-
Forest	Percentage of the 1km radius from Lat-Long position that is covered in forest (based on 2006 NLCD - natural land cover database - imagery)	-
SSG	Percentage of the 1km radius from Lat-Long GPS position that is covered in shrub scrub grassland (based on 2006 NLCD - natural land cover database - imagery)	-
Area	Area of the pond as assessed by GPS	Natural log
Veg_s	Percentage of the shore which had vegetation	-
OpenW	Percentage of the pond that is open water	Natural log
Cond	Conductivity (uS/cm) measured using Yellow Springs Instruments 556 Multi Probe System	Natural log
TDS	Total dissolved solids (g/l) measured using Yellow Springs Instruments 556 Multi Probe System	Natural log
DOmg	Dissolved oxygen (mg/L) measured using Yellow Springs Instruments 556 Multi Probe System	Natural log
Amph_Rich	Amphibian species richness at the wetland	Square root
Snail_Rich	Snail species richness at the wetland	-
Amph_MDS1	Primary axis scores of the local amphibian community NMDS ordination, based on abundances of amphibian larvae	-
Amph_MDS2	Secondary axis scores of the local amphibian community NMDS ordination, based on abundances of amphibian larvae	-
Snails_MDS1	Primary axis scores of the local snail community NMDS ordination, based on abundances	-
Snails_MDS2	Secondary axis scores of the local snail community NMDS ordination, based on abundances	-
Amph_RA1	Primary axis scores of the local amphibian community RDA ordination, based on occupancy of amphibian species	-
Amph_RA2	Secondary axis scores of the local amphibian community RDA ordination, based on occupancy of amphibian species	-
Snails_RA1	Primary axis scores of the local amphibian community RDA ordination, based on occupancy of snail species	-
Snails_RA2	Secondary axis scores of the local amphibian community RDA ordination, based on occupancy of snail species	-

5.3.3 Detection error-corrected elements of metacommunity structure

We consider symbiont metacommunities in which wetlands represent local habitats. We restricted the data set to include habitats from which we sampled at least 8 *P. regilla*. Thus, the within-host symbiont communities of these necropsied individuals are integrated to a within-wetland symbiont community composition (pres/abs occupancy for each symbiont species). In order to characterize the structure of these symbiont metacommunities and assess environmental covariate effects on symbiont composition, we use a recently developed extension of the elements of metacommunity structure framework – detection error-corrected elements of metacommunity structure (DECEMS; Mihaljevic et al. *in revision*). This statistical method relies on building a multi-species occupancy model with each species' probability of wetland-level occupancy predicted by a number of environmental covariates. We use the most supported model to predict each species' site-level occupancy and generate estimated metacommunity structures. Importantly, the model is able to estimate species-level covariate effects, making it possible to attribute metacommunity-wide pattern shifts to species-level effects.

Here we rely on a novel multi-species occupancy model derived from recent statistical advances. Jackson et al. (2012) introduced a hierarchical modeling framework that outperforms many standard methods used to identify environmental drivers of species composition, such as canonical correspondence analysis (CCA) and nonmetric multidimensional scaling (NMDS). This statistical framework uses multilevel modeling to estimate species-level random and fixed covariate effects to determine the relative contribution of environmental covariates to changing composition across space. Here we propose a simple extension to Jackson and colleague's work: correcting for detection error. Our model is structured as follows:

$$y_q \sim Binom(z_q * p_{spp[q]}, j_q)$$

$$z_q \sim Bern(\psi_q)$$

$$\psi_q = logit^{-1}(\alpha_{spp[q]} + \mathrm{B}_{spp[q]} * X_{site[q]})$$

$$\alpha_{spp[q]} \sim N(\mu_\alpha, \sigma^2_{intercept})$$

$$\beta_{spp[q]} \sim N(\mu_\beta, \sigma^2_{slope})$$

Here y_q is a vector of the number of times each symbiont species is observed at each site over j surveys ($q = 1, \ldots, nm$, where n is the number of species and m is the number of sites). In this case, sites are wetlands and the number of surveys is the number of necropsied *P. regilla*. $p_{spp[q]}$ represents the species-specific probability of detection when the species is present at the site. In practice, this value is the estimated average prevalence of a symbiont within a site (i.e. among the local *P. regilla* population). z_q represents the 'true' occurrence of the species, a Bernoulli random variable with probability, ψ_q, the species-specific probability of occurrence at a site. Via the logit transformation, ψ_q is then linearly related to a matrix of covariates, $X_{site[q]}$, with a vector of species-specific slopes, $\mathrm{B}_{spp[q]}$, and species-specific intercepts (i.e. baseline occurrences), $\alpha_{spp[q]}$. Each intercept, $\alpha_{spp[q]}$, and each slope, $\beta_{spp[q]}$, is distributed normally with means, μ_α and μ_β, and variances, $\sigma^2_{intercept}$ and σ^2_{slope}, respectively. Thus the model is able to generate bias-corrected, species-specific estimates of covariate effects by accounting for detection error.

The objective of our modeling routine is to determine whether symbiont species respond coherently to any environmental gradients, in which case we expect community composition to shift along such gradients. Concretely, we are interested in identifying covariates with significant random effects; in other words σ^2_{slope} for the covariate is significantly greater than zero. This

would demonstrate that many symbiont species respond to the covariate, but with variable effects. We are also interested in identifying covariates that have significant fixed effects; in other words μ_β for the covariate differs significantly from zero. This would demonstrate that the covariate has significant effects on occurrence that are consistent across many symbiont species.

We ran these models using a Bayesian approach to inference and parameter estimation, relying on Markov chain Monte Carlo (MCMC) sampling in which we iteratively sample from the posterior distribution of each parameter. All models were run in the open-source software JAGS (Just Another Gibbs Sampler; http://mcmc-jags.sourceforge.net/) via the open-source statistical software, *R* (R Core Team 2014). We used uninformative priors and ran each model with three MCMC chains. Models were run with an initial 80,000 iteration adaptation phase, followed by an 80,000 iteration burn-in period, and then 1,500 iterations were stored, thinning by 50 iterations, for a total of 235,000 iterations per model run. We assessed convergence using the potential scale reduction factor, $\hat{R}$ (Gelman 1996).

We created a model for each year individually, and then a global model for all years' data combined. Within each year (and then again for the global model), we removed any collinear covariates and ran the model with all remaining covariates, which we call the full model (Table 5.3). From the full model run, we estimated the 95% highest density intervals (HDI) of the slope parameters' posterior distributions. In the subsequent model run, which we call the reduced model, we only included covariates whose 95% HDI σ^2_{slope} or μ_β did not include zero. Based on this reduced model, we then conducted model selection using an information theoretic approach. Specifically, for each nested model – including a null model that only included the random intercepts – we calculated the Watanabe-Akaike information criteria (WAIC). WAIC, which is analogous to AIC, has several advantages over other Bayesian-type information criteria, and is

the preferred metric for occupancy models (Gelman et al. 2013, Watanabe 2013, Hooten and Hobbs 2014).

Table 5.3. List of collinear covariates in each year and the covariates that we removed, accordingly. In 2012, because there was such a high level of collinearity, and so many covariates had to be removed, we reported which covariates were included in the full model.

2009	**2010**
Lat - Long - Elev	Lat - Long - Elev
Forest - SSG	Forest - SSG
Cond - TDS	Cond - TDS
Amph_RA1 - Amph_MDS1	Slope - Area
Amph_Rich - Amph_RA2	Snails_RA1 - Snails_MDS1 - Snail_Rich
REMOVED	REMOVED
Long, Elev, SSG, TDS, Amph_MDS1, Amph_RA2	Long, Elev, SSG, TDS, Slope, Snails_MDS1, Snails_RA1
2011	**2012**
Lat - Long - Elev	Lat - Long - Elev - Aspect - Area - TDS - Snails_MDS2 - Amph_RA1 - Amph_RA2 - Slope - Forest - SSG - Amph_rich - Amph_MDS2 - Snails_RA2
Forest - SSG	
Cond - TDS	
OpenW - Domg	
Snails_RA1 - Snails_MDS2 - Snail_Rich	
Amph_RA1 - Amph_MDS2	
REMOVED	*INCLUDED*
Long, Elev, SSG, TDS, OpenW, Snails_MDS2, Snails_RA1, Amph_MDS2	Aspect, Hydro, OpenW, Domg, Amph_Rich, Amph_MDS1, Snails_RA2

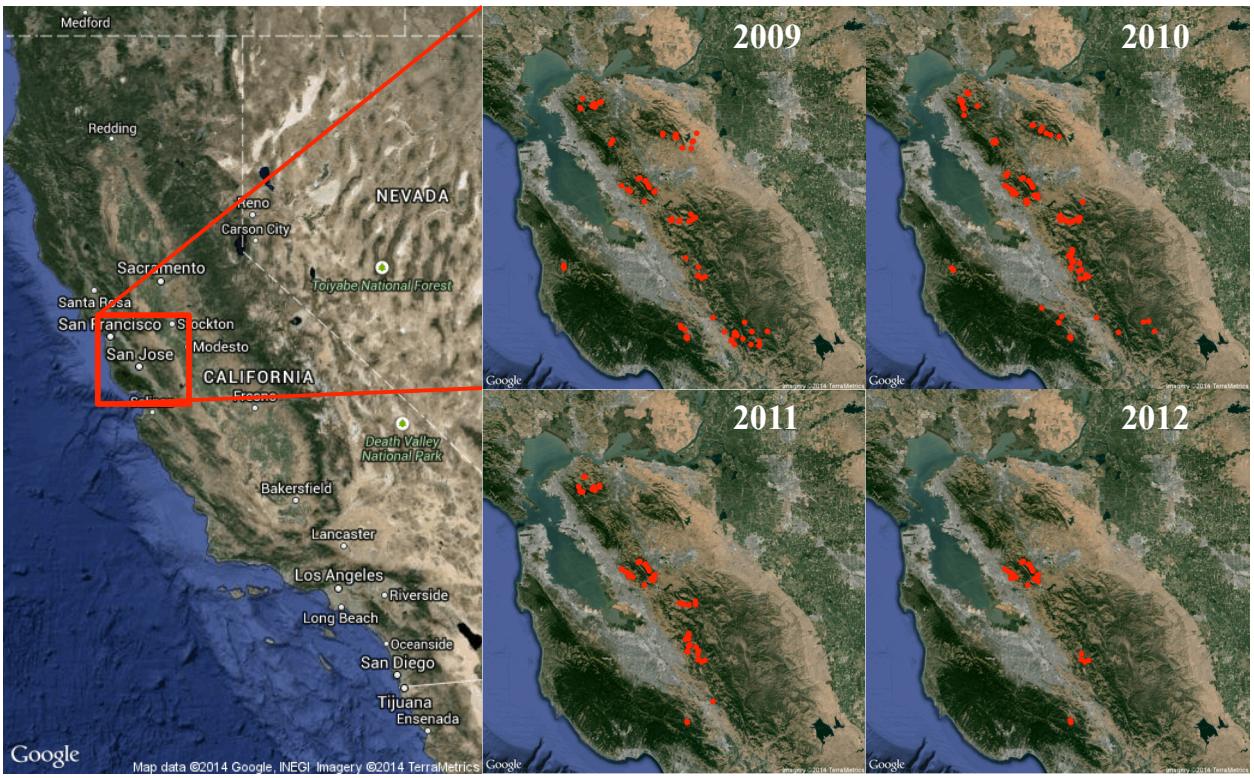
From the most supported model, we compiled posterior estimates of z_q into posterior site-by-species matrices, $\boldsymbol{Z_{post}}$. We created 500 $\boldsymbol{Z_{post}}$ for each year, and for each matrix we calculated the three elements of metacommunity structure – coherence, turnover, and boundary clumping. Based on these metrics, we assigned a categorical metacommunity structure to each $\boldsymbol{Z_{post}}$. Thus, from the most supported multi-species occupancy model, we are able to generate pseudo-posterior estimates of each element of metacommunity and an overall probability distribution of metacommunity structure. We generated the elements of metacommunity structure using the package '*metacom'* in R (Dallas 2014). The details of calculating these metrics have been reviewed extensively elsewhere (e.g. Leibold and Mikkelson 2002, Presley et al. 2010, Mihaljevic et al. *in revision*).

5.4 Results

5.4.1 Surveys

From the data set we included a total of 266 surveys of 158 wetlands over 4 years (2009-2012), with necropsy data from 2887 metamorphic *P. regilla* (Figure 5.1). Across all wetlands and years we commonly encountered 8 symbiont species of *P. regilla* that occur in the aquatic environment, 6 of which were trematodes with complex life cycles and 2 of which were intestinal protists that are directly transmitted among amphibians (Table 5.1). Other symbiont species were encountered, but these symbionts are acquired in the terrestrial life stage of the amphibian. Thus, any metamorphs with these symbionts were not included in the study, as we could not be certain that these metamorphs originated at the wetland of interest. Two other symbionts were not included in our study because they were observed at less than 8% of wetlands within a given year and in less than 6% of all surveys over the 4 year period. There were larvae from up to six amphibian species present at each site: *Taricha torosa* (rough-skinned newt), *Ambystoma californiense* (California tiger salamander), *Lithobates catesbeianus* (American bullfrog), *Anaxyrus boreas* (Western toad), *Rana draytonii* (California red-legged frog), and our focal host *Pseudacris regilla* (Pacific chorus frog). We also encountered up to five snail species at each site: *Radix* spp., *Lymnaea* spp., *Helisoma trivolvis*, *Gyraulus* spp., and *Physa* spp.

Figure 5.1. Map of wetland sites sampled in the San Francisco Bay Area of California, USA from 2009-2012.



5.4.2 Symbiont composition over time

In 2009, we detected 6 common symbionts (Alar, Glob, Echi, Mano, Opal, and Rib) across 77 wetlands and 964 necropsied *P. regilla*. From the full covariate model, the reduced model included Aspect, Amph_MDS2, Snails_MDS1, Snails_RA1, and Snails_RA2. However, the most supported models, based on WAIC selection, contained various combinations of only Aspect + Amph_MDS2 + Snails_RA2 + Snails_MDS1 (Table 5.4), with a significant random effect of Snails_RA2 (95% HDI σ_{slope}: 0.02 – 1.90). We used the model containing Aspect + Amph_MDS2 + Snails_RA2 to generate the 500 $\boldsymbol{Z_{post}}$. Approximately 45% of these posterior metacommunities exhibited significant nested or quasi-nested structures; however ~50% of the matrices were randomly structured (Figure 2a). While *Opalina* spp. and *Echinostoma* spp. were present at most sites, the other species' ranges were more restricted. The model estimated that

Ribeiroia ondatrae was positively affected by Snails_RA2, *Cephalogonimus* spp. (Glob) was positively affected by Amph_MDS2, and *Alaria* spp. was negatively affected by Aspect (Figure 5.3).

Table 5.4. The top five candidate models and associated WAIC and ΔWAIC values for each year and for the combined data set. The WAIC values of the full, reduced, and null models are added for reference. Note that in 2012, the null model was included in the top five candidate models.

2009			**2010**		
Model	WAIC	ΔWAIC	**Model**	WAIC	ΔWAIC
Aspect + Amph_MDS2 + Snails_RA2	1454.78	-	Lat + Snails_MDS2 + Snails_RA2	2195.75	-
Amph_MDS2 + Snails_RA2	1456.36	1.58	Lat + Snails_RA2	2195.98	0.23
Aspect + Amph_MDS2 + Snails_RA2 + Snails_MDS1	1456.44	1.66	Lat	2196.57	0.83
Aspect + Snails_RA2 + Snails_MDS1	1456.60	1.82	Lat + Snails_MDS2 + Snails_RA2 + Amph_RA2	2197.26	1.51
Aspect	1457.33	2.55	Snails_MDS2 + Snails_RA2 + Amph_RA2	2197.47	1.72
FULL	1486.24	31.46	FULL	2224.97	29.23
REDUCED	1459.75	4.97	REDUCED	2205.51	9.77
NULL	1458.63	3.85	NULL	2199.98	4.24
2011			**2012**		
Model	WAIC	ΔWAIC	**Model**	WAIC	ΔWAIC
Amph_RA1 + Amph_RA2	1636.45	-	Hydro	491.75	-
Amph_RA2	1636.66	0.22	Amph_Rich	492.02	0.28
Aspect + Amph_RA2	1636.74	0.30	*NULL*	492.32	0.58
Aspect	1636.95	0.51	Hydro + Amph_Rich	492.87	1.12
Aspect + Amph_RA1 + Amph_RA2	1637.41	0.97	Hydro + OpenW	494.05	2.30
FULL	1735.61	99.17	FULL	517.04	25.30
REDUCED	1656.54	20.10	REDUCED	498.77	7.02
NULL	1638.88	2.43	-	-	-
All Years					
Model	WAIC	ΔWAIC			
Amph_RA1 + Amph_RA2	6320.57	-			
Amph_RA1 + Amph_RA2 + Snails_RA2 + Veg_s	6320.93	0.37			
Amph_RA1 + Amph_RA2 + Veg_s	6321.98	1.41			
Amph_RA1 + Veg_s	6322.18	1.61			
Amph_RA1 + Snails_RA2 + Veg_s	6322.28	1.71			
FULL	6338.86	18.30			
REDUCED	6328.37	7.80			
NULL	6323.25	2.68			

Figure 5.2. Estimated metacommunity structures for each year. These are ordinated matrices, based on each species' wetland-level the probability of occurrence, as estimated by the most supported model. The gray scale within the ordinated matrix corresponds to probability of occurrence, as shown in the upper-left legend. To the side of each ordinated matrix is the distribution of structures estimated for each of 500 $\boldsymbol{Z}_{post}$. Species labels are on the x-axis and each wetland is a row on along the y-axis of the ordinated matrices.

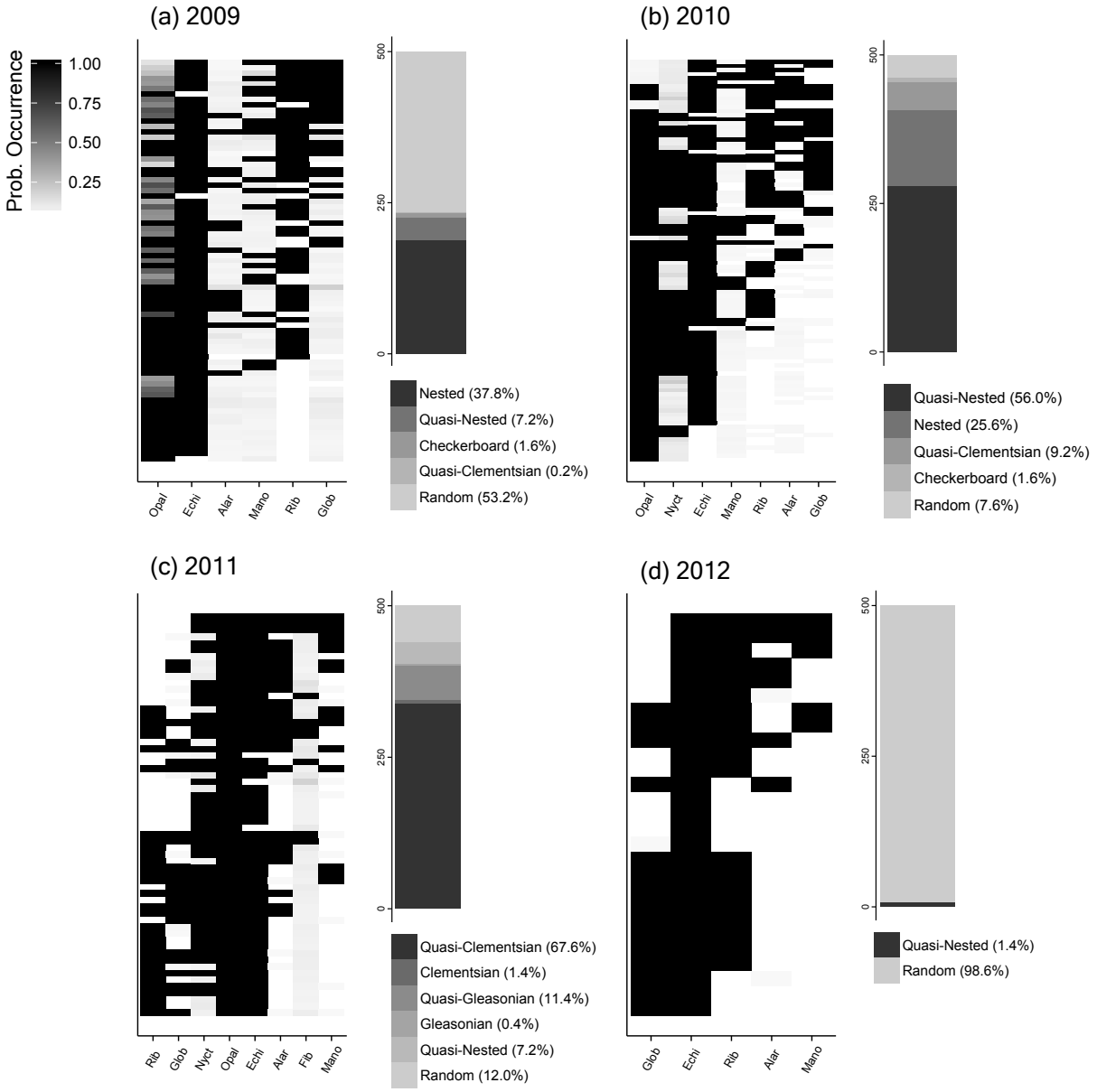
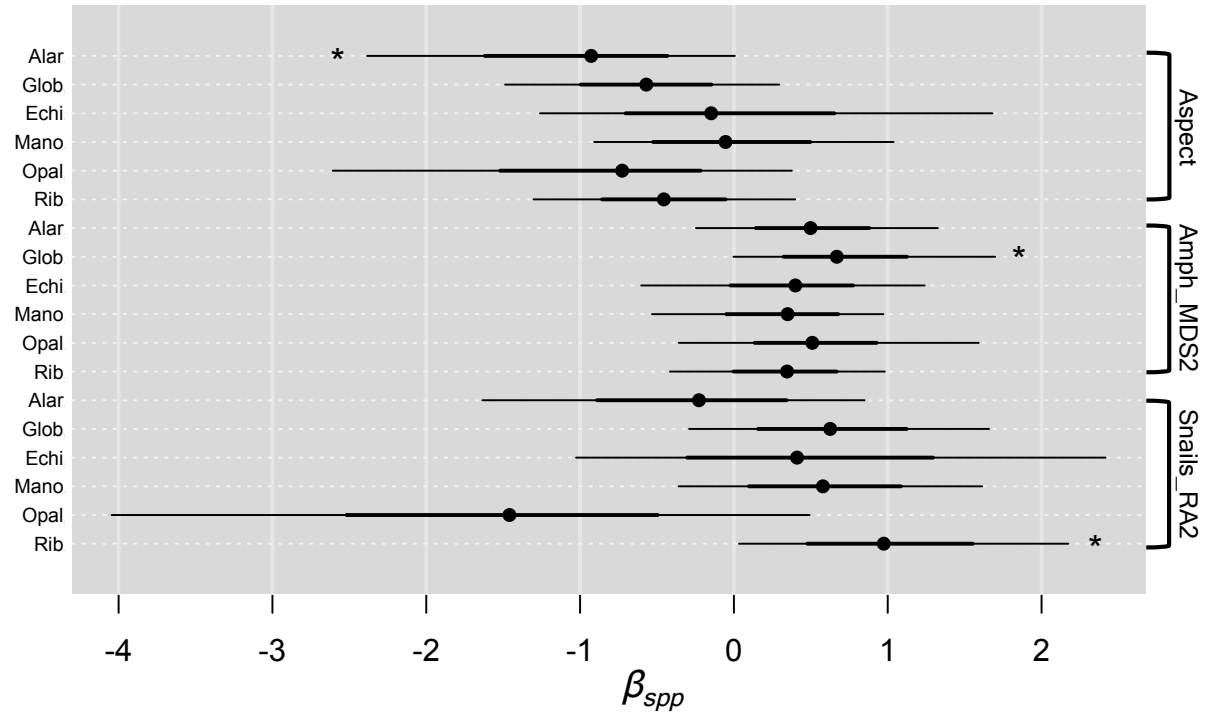


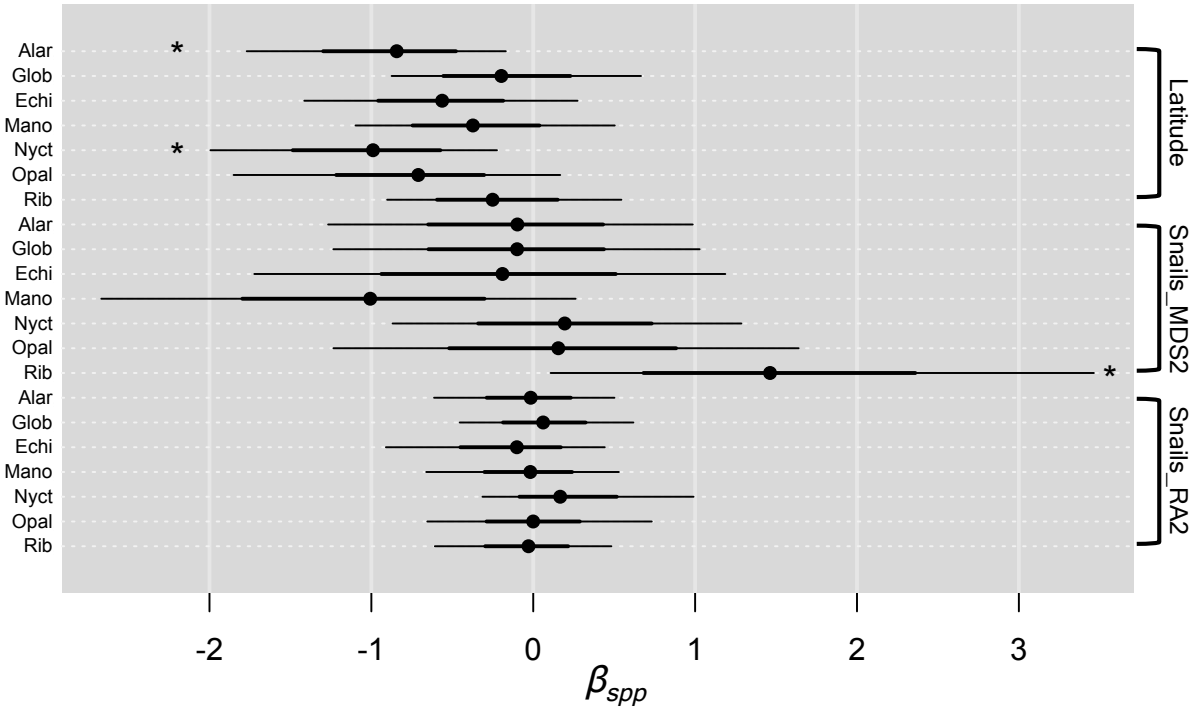
Figure 5.3. Estimates of species-specific covariate effects in 2009. Black dots represent median estimates from the posterior probability, while dark horizontal bars represent 68% HDI, and lighter horizontal bars represent 95% HDI. Stars mark intervals for which zero is not included in the 95% HDI (i.e. significant species-level effects).



In 2010, we encountered 7 common symbionts across 98 wetlands and 1011 sampled *P. regilla*, adding *Nyctotherus* spp., another protozoan, to those encountered in 2009. In this year, the reduced model included Latitude, Snail_Rich, Snails_MDS2, Amph_RA2, and Snails_RA2. The most supported models, however, only included combinations of Latitude, Snails_MDS2, Snails_RA2 and Amph_RA2 (Table 5.4). We found a significant negative, fixed effect of Latitude (95% HDI μ_β: -0.95 – -0.18) and a significant random effect of Snails_MDS2 (95% HDI σ_{slope}: 0.15 – 1.68) showing that most species tended to have lower probabilities of occupancy at high latitude wetlands, and, again, compositional change was driven by differences in snail intermediate host community composition. Using the Latitude + Snails_MDS2 + Snails_RA2 model to generate the $\boldsymbol{Z_{post}}$ matrices, the symbiont metacommunity exhibited quasi-

nested or nested structure with ~80% probability (Figure 5.2b). *Opalina* spp., *Nyctotherus* spp., and *Echinostoma* spp. were very common, with the other symbionts distributions nested within these ranges. *R. ondatrae* and *Manodistomum syntomentera* had positive and negative responses to Snails_MDS2, respectively, while *Alaria* spp. and *Nyctotherus* spp. both had negative effects of Latitude (Figure 5.4).

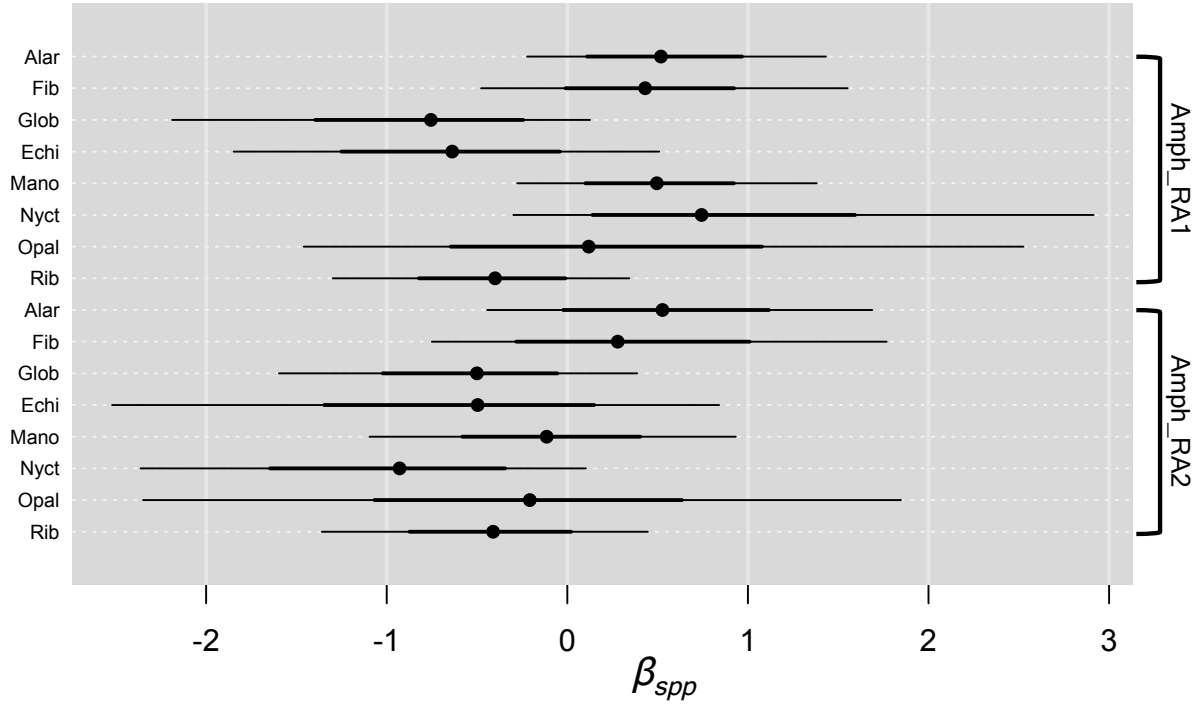
Figure 5.4. Estimates of species-specific covariate effects in 2010. Figure is laid out as in Figure 3.



In 2011, we encountered the full set of 8 symbiont species across 61 wetlands and 607 sampled *P. regilla*. Latitude, Aspect, Forest, Veg_s, Snail_Rich, Amph_RA1, Amph_RA2 and Snails_RA2 were all included in the reduced model. However, only combinations of Amph_RA1, Amph_RA2 and Aspect were included in the most supported models (Table 5.4). There was a significant random effect of Amph_RA1 (95% HDI σ_{slope}: 0.14 – 1.50), showing

that, in contrast to the previous two years, symbiont compositional change was driven by the composition of the alternative amphibian host composition across wetlands. Concordantly, the symbiont metacommunity structure in this year was noticeably different from the previous two years, with the $\boldsymbol{Z}_{post}$ demonstrating mostly quasi-Clementsian or Clementsian structure (~70%). Our most supported model estimated that *R. ondatrae* and *Cephalogonmius* spp. respond predominantly negatively to the Amph_RA1 axis, while *Alaria* spp., *Fibricola* spp., and *M. syntomentera* have negative responses (Figure 5.5). Thus, the symbiont community responded coherently to the amphibian host community composition.

Figure 5.5. Estimates of species-specific covariate effects in 2011. Figure is laid out as in Figure 3.



In 2012 the sample size of wetlands was much lower (n=30) with 305 *P. regilla* sampled and only 5 commonly encountered symbionts: Glob, Echi, Rib, Mano, Alar. There was a notably high degree of collinearity among covariates in this year as well (Table 5.4). The reduced

occupancy model included only Aspect, OpenW, Amph_Rich, and Hydro. While the most supported models contained Hydro, OpenW and Amph_Rich, these could not be distinguished from the null model that included only the random intercept (i.e. species-level baseline occurrence probability) (Table 5.4). In support of this result, the metacommunity structure was random (98.6%).

When all the data was pooled, the reduced model included Conductivity, Veg_s, Amph_MDS1, Amph_RA1, Amph_RA2, Snail_Rich, and Snails_RA2. However, the best models only included combinations of Amph_RA1, Amph_RA2, Snails_RA2 and Veg_s (Table 5.4), with a significant random effect of Amph_RA1 (95% HDI σ_{slope}: 0.28 – 0.79) and Amph_RA2 (95% HDI σ_{slope}: 0.20 – 0.73). The metacommunity structure of the full data set showed an even split between quasi-Clementsian and quasi-nested structure (45% and 32%, respectively; Figure 5.6). Our model estimated a positive effect of Amph_RA1 on *Fibricola* spp. and *M. syntomentera* occupancy and a negative effect on *R. ondatrae* and *Cephalogonmius* spp. (Figure 5.7).

Figure 5.6. Symbiont metacommunity structure with all data combined.

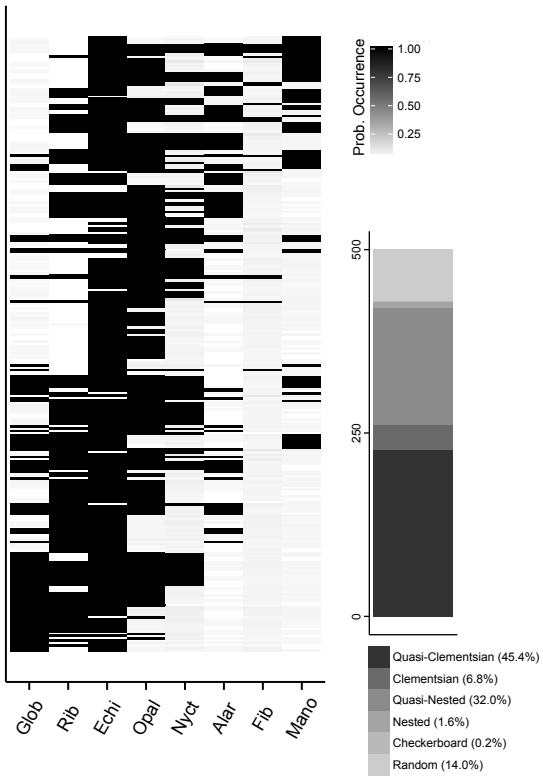
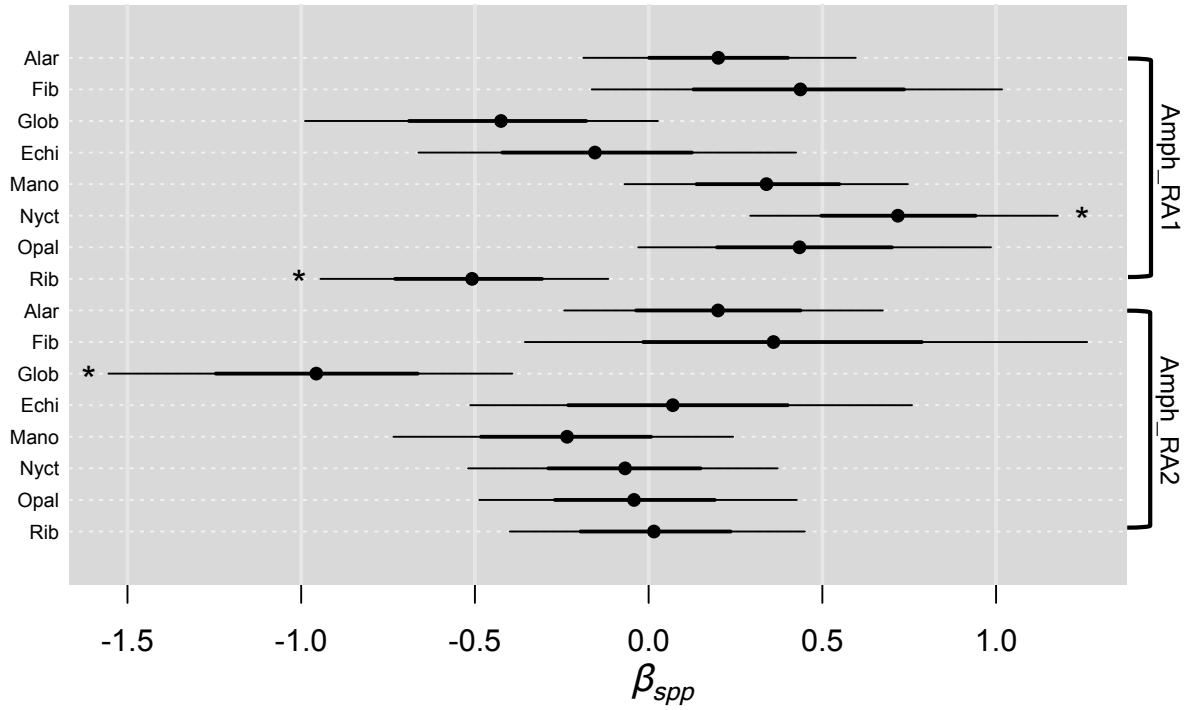
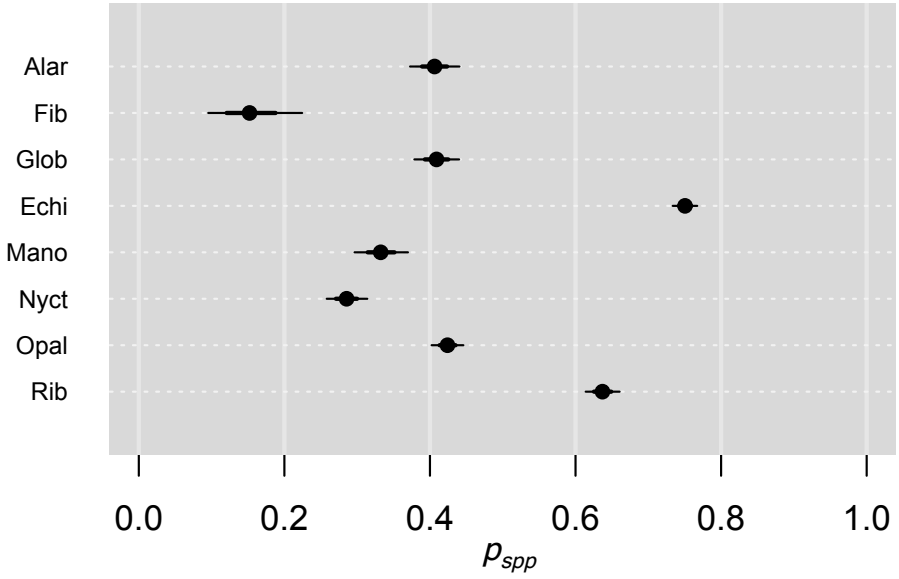


Figure 5.7. Estimates of species-specific covariate effects for all data combined. Figure is laid out as in Figure 3.



Our modeling technique was also able to estimate species-level probabilities of detection, which, in light of our sampling design, are equivalent to each species' average prevalence at a wetland when it occurs. Each symbiont's prevalence was rather consistent among years, and the model with all years' data was able to very precisely estimate this prevalence for each species (Figure 5.8).

Figure 5.8. Estimates of species-specific detection probabilities using the combined data set.



5.5 Discussion

Although metacommunity theory is still developing to accommodate host-symbiont dynamics, metacommunity tools offer informative ways to understand how and why symbiont community changes through space and time. In this study we applied newly developed metacommunity methods to explore how the symbiont metacommunity structure of the Pacific chorus frog, *P. regilla*, changes across a four-year period. We found that in most years, the symbiont community exhibited discernable metacommunity structures, although the support for these structures was not always strong. However, the structures were not consistent from year to year. The most supported models of symbiont species composition among years typically included aspects of first intermediate host (snail) composition or alternative secondary intermediate host (amphibian) composition, as well as geographical wetland characteristics, such as latitude and aspect.

We found that the symbiont metacommunity structure alternated between random, quasi-nested, and quasi-Clementsian among years. Because our modeling methods are able to estimate species-specific responses to covariates, we are able to directly link these overall metacommunity patterns to species-level effects. In 2009, when metacommunity structure had a high probability of being random, our model predicted that different symbiont species were responding to different environmental gradients. Our model estimated that *Ribeiroia ondatrae* was positively affected by Snails_RA2, *Cephalogonimus* spp. (Glob) was positively affected by Amph_MDS2, and *Alaria* spp. was negatively affected by Aspect. Indeed, when a metacommunity exhibits random structure, it is supposed that species are not coherently responding to a common environmental gradient (Leibold and Mikkelson 2002b, Presley et al. 2010a). Here our model is able to directly support this supposition. The nested and quasi-nested structures observed in some $\boldsymbol{Z_{post}}$ in 2009 were likely driven by the fact that *Echinostoma* spp. and *Opalina* spp. were observed (or predicted to occur) at the majority all sites. This makes sense because these Echinostomes and have rather broad intermediate host ranges with vagile definitive hosts (Roberts et al. 1996). *Opalina* spp. has a direct transmission route and uses amphibians as its only host (and amphibians were present at every surveyed wetland) (Smyth and Smyth 1980). The other symbiont species had lower overall prevalence across the study area. Thus, the nested patterns observed in some $\boldsymbol{Z_{post}}$ might statistical artifacts, rather than a true demonstration of coherent species responses to a common gradient (Fischer and Lindenmayer 2002).

In 2010, the symbiont metacommunity had an ~80% probability of having quasi-nested or nested structure, however an ~20% chance of being randomly structured. The dominant structure is quasi-nested because although the distributions of *R. ondatrae*, *M. syntomentera*,

Alaria spp., and *Cephalogonimus* spp. were nested within those of *Opalina* spp., *Echinostoma* spp., and *Nyctotherus* spp., there were many instances were the former 4 species did not occur with one another to create a fully nested metacommunity. This is likely due to the fact that these species were again responding to unique environmental gradients. Indeed, *R. ondatrae* and *M. syntomentera* had positive and negative responses to Snails_MDS2, respectively, while *Alaria* spp. and *Nyctotherus* spp. both had negative effects of Latitude. These divergent coherent responses to snail composition and latitude likely drive the ambiguity in structure classification.

In 2011, the metacommunity showed predominantly quasi-Clementsian structure. Clementsian structure demonstrates that species composition shifts along an environmental gradient, forming discrete and unique species groups along the gradient (Leibold and Mikkelson 2002b). The "quasi" nature of the structure is likely due to a relatively small symbiont community and many embedded absences (Gotelli 2000, Presley et al. 2010a). This structure seems to be driven by a shift from communities that include the common core species – *Opalina* spp., *Echinostoma* spp., and *Nyctotherus* spp. – as well as *R. ondatrae* and *Cephalogonmius* spp. to communities with the core species and *Alaria* spp., *Fibricola* spp. Our model estimated that *R. ondatrae* and *Cephalogonmius* spp. respond predominantly negatively to the Amph_RA1 axis, while *Alaria* spp., *Fibricola* spp., and *M. syntomentera* have negative responses. This indicates that the Clementsian structure is driven by a coherent response to a gradient of amphibian host species composition at the sites. It is likely that in 2012 that our smaller sample size and the presence of fewer symbiont species hindered our ability to discern metacommunity structure and find a model, other than the null, that best explained community composition. The smaller number of wetlands and species leads to a smaller and sparser matrix, which results in less power to detect patterns (Gotelli and Graves 1996, Ulrich and Gotelli 2013).

Similarly to 2011, when all the data was combined, the metacommunity showed quasi-Clementsian structure, with two community types distinguished by the presence of the core species, *R. ondatrae*, and *Cephalogonmius* spp., or the core species, *Fibricola* spp. and *M. syntomentera*. Indeed our model estimated a positive effect of Amph_RA1 on *Fibricola* spp. and *M. syntomentera* occupancy and a negative effect on *R. ondatrae* and *Cephalogonmius* spp. Thus, across the whole data set, a coherent response to the amphibian secondary host community composition drives metacommunity structure.

Overall, across the years, compositional shifts were driven predominantly by local differences in snail and amphibian host community compositions. It makes sense that differences in local snail host species richness and relative abundances affect community composition. The majority of the symbionts are trematodes that use snails as their first intermediate host, however, some of the symbionts are much more host-specific at this life-stage than others. For instance, while *Echinostoma* spp. can use four out of the five snail species as intermediate hosts, all of the other trematode species must use *H. trivolvis* (Prudhoe and Bray 1982, Roberts et al. 1996, Johnson et al. 2004).Thus, the presence of multiple snail species may result in competitive dynamics that alter the abundance of *H. trivolvis* (Brown 1982), leading to changes in host availability that could differentially affect the different symbiont species, depending on their ability to compete for snail hosts.

The local amphibian community composition was also a main driver of symbiont community change across space. This effect could be driven by influences of host composition on colonization and transmission dynamics in the wetlands. For instance, field surveys have shown that the richness of the symbiont community in the wetland correlates with the richness of host and non-host (e.g. insects) taxa at the wetland (Johnson et al. 2013a). This is likely an affect

of enhanced colonization opportunities, in that wetlands that harbor high biodiversity might attract many definitive hosts that deposit trematode eggs. Positive relationships between host diversity and symbiont diversity are seen across a wide variety of taxa, including trematodes and bird hosts (Hechinger and Lafferty 2005, Pedersen et al. 2005). However, host richness within a wetland is also negatively associated with the realized transmission of many of the trematodes encountered in this study (Johnson et al. 2013a, 2013b). This is due to the fact that amphibian host species differ in their host competence, whereby some hosts can rapidly clear infections, while others can become heavily infected. Furthermore, higher symbiont diversity within a wetland correlates with decreased *R. ondatrae* infection in *P. regilla*, which is likely due to more prominent intra-host dynamics with increased symbiont richness (Johnson et al. 2013a).

In the first three years of the study, wetland aspect and latitude were also significant drivers of symbiont composition. In 2009, although aspect did not have a significant fixed or random effect in the model, there seemed to be a trend where many symbiont species responded negatively to aspect. This means that occupancy was less probable at wetlands with southerly or westerly facing slopes, which tend to receive higher solar radiation and are warmer, drier habitats in general. It seems likely that more heat-sheltered, easterly facing wetlands would be more hospitable to definitive hosts of many of these trematodes. Also, in these oak chaparral habitats adult amphibians tend to aggregate more on easterly facing slopes, which are cooler and have more shrub cover (Block and Morrison 1998). In 2010, there was a significant negative effect of latitude. In our sampling design latitude is strongly correlated with longitude, which is in turn correlated with elevation. Therefore, in general, symbiont species tended to have higher probabilities of occurrence at the southeastern, higher elevation wetlands. The southeastern wetlands tend to be situated in larger areas of undeveloped land, whereas the northwestern

wetlands are bordered by water and are situated in a narrower strip of land (Figure 5.1). Based on a simple species-area relationship (Rosenzweig 1995), one might expect that these southeastern areas have a higher diversity of vertebrate species , which might lead to overall higher occupancy probabilities for the symbionts.

5.6 Conclusion

Here we saw that the *P. regilla* symbiont metacommunity structure across a large region of wetlands was inconsistent over a four-year sampling period, at times showing random, nested, or Clementsian structures. Local symbiont composition typically responded strongest to spatial variation in snail and amphibian host community composition, as well as geographic characteristics, namely, latitude and aspect. The apparent lack of metacommunity structure in some years was typically a result of non-coherent species-level responses to different environmental gradients. Inter-annual variation in metacommunity structure was likely impacted by inter-annual variability in the composition of sampled symbionts from year to year, variation in sampling breadth across space, and inter-annual changes to local host community compositions.

Our results illustrate two important aspects of this system's ecology. First, the symbionts in this system do not typically respond to the same environmental gradients, which emphasizes that predictions about symbiont community composition will have to integrate across the factors that influence species-level responses. This result further justifies the use of multi-species occupancy modeling in understanding metacommunity dynamics, as these models can integrate species-level responses to multiple covariates. Second, our results suggest that, while the local aquatic intermediate host composition certainly impacts symbiont occupancy patterns, more fine-

scale data on definitive host resource use will be necessary to better understand how colonization opportunities affect local and regional occupancy patterns.

It is likely that our results are relevant to other systems of symbionts that include species with complex life cycles. These symbionts rely on multiple host species that often utilize vastly different environments. Predicting the symbiont composition in any one environment will require integrating the effects of symbiont colonization and transmission across environment types. Thus, further understanding symbiont community dynamics across space and time will require a synthesis of ecological and evolutionary dynamics that occur at multiple spatial scales. An emphasis on building analytical tools and methods that link dynamics across scales should be a priority in this field.

CONCLUSION

In my first chapter I showed that co-exposure to multiple *Ranavirus* species, viruses of serious conservation concern, can enhance viral infectivity, lead to larger epidemics, and increase the hosts' probability of mortality. By conducting two experiments at the host-scale and then at the population-scale, I was better able to demonstrate and understand the mechanisms leading to an effect of co-exposure. This research has direct implications for the conservation of amphibian populations, globally. In future studies, I would like to conduct immunological assays to determine the intra-host mechanisms that result in higher FV3 infectivity when ATV co-occurs. I also plan to sequence viral DNA from field-collected specimen in our California field sites (see Chapter 5) in order to determine if FV3- and ATV-like viruses co-occur in the region. Finally, I would like to scale up the system to understand how host biodiversity interacts with viral diversity to affect *Ranavirus* transmission dynamics.

For Chapter 2, I built a multi-host species epidemiological model to understand how pathogen transmission type and host community composition affects how host biodiversity influences transmission in local host communities. This model was able to formally test and verify mathematical and verbal predictions of diversity-disease relationships, and extend these models to generate more specific predictions about where, when, and how host biodiversity should affect disease transmission. Importantly, we showed that high variability in biodiversity's effect on transmission could hinder finding generalizable trends (e.g. biodiversity should always decrease disease transmission). In future studies, I would like to extend the model to include multiple pathogens, which would complement my empirical work. I also plan to put the model into a spatial context to look at how host migration among local communities in a metacommunity will affect regional transmission dynamics, including local pathogen

colonization-extinction dynamics.

My last three chapters were focused on applying metacommunity theory and analytical tools to the study of host-symbiont (including pathogen) interactions. I first reviewed how metacommunity theory could help answer many outstanding questions in the field, highlighting that symbiont communities constitute metacommunities at multiple spatial scales, and that these systems could also be used to test metacommunity theory predictions. I then built upon existing metacommunity analytical tools to incorporate the benefits of occupancy modeling, which can correct for detection error. Detection error is likely a large issue in sampling symbiont communities due to their typically small size and cryptic nature. Finally, I implement my newly developed analytical framework to study how amphibian parasite metacommunities are structured across four years of sampling. The methods I employed allowed to estimate species-specific responses to environmental covariates. This approach should be very useful to understanding how parasites of particular concern are distributed across the landscape in the context of other symbiont community members. I plan to apply these methods to data sets of microbial communities among hosts, in an effort to understand how microbial community composition changes in relation to host traits and regional transmission dynamics.

With this dissertation, I have shown that both pathogen and host biodiversity can affect pathogen transmission in complex ways, and that studying these dynamics at multiple spatial scales is integral to gain a more complete understanding. Future research would benefit from understanding how pathogen and host diversity interact to affect disease severity and pathogen transmission. This will require the integration of observation, experimental, and theoretical studies that span spatial scales. However, by accomplishing this goal, we will move toward a more complete and predictive framework for the study of disease ecology in general.

REFERENCES

Abrams, A. J., D. C. Cannatella, D. M. Hillis, and S. L. Sawyer. 2013. Recent host-shifts in ranaviruses: signatures of positive selection in the viral genome. The Journal of general virology 94:2082–93.

Alizon, S., and M. van Baalen. 2008. Multiple infections, immune dynamics, and the evolution of virulence. The American naturalist 172:E150–68.

Alizon, S., a Hurford, N. Mideo, and M. Van Baalen. 2009. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. Journal of evolutionary biology 22:245–59.

Alizon, S., and S. Lion. 2011. Within-host parasite cooperation and the evolution of virulence. Proceedings. Biological sciences / The Royal Society.

Alizon, S., F. Luciani, and R. R. Regoes. 2011. Epidemiological and clinical consequences of within-host evolution. Trends in Microbiology 19:24–32.

Alizon, S., J. C. de Roode, and Y. Michalakis. 2013. Multiple infections and the evolution of virulence. Ecology letters:1–12.

Allen, L., V. Brown, C. Jonsson, S. Klein, S. Laverty, K. Magwedere, J. Owen, and P. van den Driessche. 2012. Mathematical modeling of viral zoonoses in wildlife. Natural resource modeling 25:5–51.

Anderson, R. M., R. M. May, and B. Anderson. 1992. Infectious Diseases of Humans: Dynamics and Control (Oxford Science Publications). Page 768. Oxford University Press, USA.

Balmer, O., and M. Tanner. 2011. Prevalence and implications of multiple-strain infections. The Lancet infectious diseases 11:868–78.

Bell, A. S., J. C. de Roode, D. Sim, and A. F. Read. 2006. Within-host competition in genetically diverse malaria infections: parasite virulence and competitive success. Evolution; international journal of organic evolution 60:1358–71.

Blasco-Costa, I., J. M. Waters, and R. Poulin. 2012. Swimming against the current: genetic structure, host mobility and the drift paradox in trematode parasites. Molecular ecology 21:207–217.

Block, W., and M. Morrison. 1998. Habitat Relationships of Amphibians and Reptiles in California Oak Woodlands. Journal of herpetology 32:51–60.

Bollinger, T. K., J. Mao, D. Schock, R. M. Brigham, and V. G. Chinchar. 1999. Pathology, isolation, and preliminary molecular characterization of a novel iridovirus from tiger salamanders in Saskatchewan. Journal of wildlife diseases 35:413–29.

Ter Braak, C. J. F., and I. C. Prentice. 1988. A theory of gradient analysis. Advances In Ecological Research 18:271–317.

Brown, K. 1982. Resource overlap and competition in pond snails: an experimental analysis. Ecology 63:412–422.

Brunner, J., K. Barnett, and C. Gosier. 2011. Ranavirus Infection in Die-offs of Vernal Pool Amphibians in New York , USA. Herpetological Review 42:76–79.

Brunner, J., and J. Collins. 2009. Testing assumptions of the trade-off theory of the evolution of parasite virulence. Evolutionary Ecology Research 11:1169–1188.

Brunner, J. L., K. Richards, and J. P. Collins. 2005. Dose and host characteristics influence virulence of ranavirus infections. Oecologia 144:399–406.

Burton, A., M. Sam, C. Balangtaa, and J. Brashares. 2012. Hierarchical Multi-Species Modeling of Carnivore Responses to Hunting, Habitat and Prey in a West African Protected Area. PloS one 7.

Cavender-Bares, J., K. H. Kozak, P. V. a Fine, and S. W. Kembel. 2009. The merging of community ecology and phylogenetic biology. Ecology letters 12:693–715.

Chase, J. M. 2005. Towards a really unified theory for metacommunities. Functional Ecology 19:182–186.

Chen, G., and J. Robert. 2011. Antiviral immunity in amphibians. Viruses 3:2065–86.

Chinchar, V. G., J. Robert, and A. T. Storfer. 2011. Ecology of viruses infecting ectothermic vertebrates - The impact of ranavirus infections on amphibians. Studies in Viral Ecology:231–259.

Chinchar, V., A. Hyatt, T. Miyazaki, and T. Williams. 2009. Family Iridoviridae : Poor Viral Relations No Longer. Lesser Known Large dsDNA ….

Choisy, M., and J. C. de Roode. 2010. Mixed infections and the evolution of virulence: effects of resource competition, parasite plasticity, and impaired host immunity. The American naturalist 175:E105–18.

Clements, F. E. 1916. Plant Succession: An Analysis of the Development of Vegetation. Page 512. Carnegie Institution of Washington.

Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecology letters 8:1175–82.

Cox, F. 2001. Concomitant infections, parasites and immune responses. Parasitology 122:S23–S38.

Dallas, T. 2014. metacom: an R package for the analysis of metacommunity structure. Ecography.

Dallas, T., and S. J. Presley. 2014. Relative importance of host environment, transmission potential and host phylogeny to the structure of parasite metacommunities. Oikos.

Daszak, P. 2000. Emerging Infectious Diseases of Wildlife-- Threats to Biodiversity and Human Health. Science 287:443–449.

Day, T., S. Alizon, and N. Mideo. 2011. Bridging Scales in the Evolution of Infectious Disease Life Histories: Theory. Evolution 65:3448–3461.

Dieckmann, U., J. A. J. Metz, M. W. Sabelis, and K. Sigmund. 2002. Adaptive dynamics of infectious diseases: in pursuit of virulence management. Page 532. Cambridge University Press.

Dobson, A. 2004. Population dynamics of pathogens with multiple host species. The American naturalist 164 Suppl :S64–78.

Dorazio, R. M., M. Kéry, J. A. Royle, and M. Plattner. 2010. Models for inference in dynamic metacommunity systems. Ecology 91:2466–2475.

Dove, A. D. M. 2006. Defining parasite communities is a challenge for neutral theory. The Journal of parasitology 92:673–5.

Eaton, H. E., J. Metcalf, E. Penny, V. Tcherepanov, C. Upton, and C. R. Brunetti. 2007. Comparative genomic analysis of the family Iridoviridae: re-annotating and defining the core set of iridovirus genes. Virology journal 4:11.

Economo, E. P. 2011. Biodiversity conservation in metacommunity networks: linking pattern and persistence. The American naturalist 177:E167–80.

Erős, T., P. Sály, P. Takács, C. L. Higgins, P. Bíró, and D. Schmera. 2014. Quantifying temporal variability in the metacommunity structure of stream fishes: the influence of non-native species and environmental drivers. Hydrobiologia 722:31–43.

Ezenwa, V. O., R. S. Etienne, G. Luikart, A. Beja-Pereira, and A. E. Jolles. 2010. Hidden Consequences of Living in a Wormy World: Nematode-Induced Immune Suppression Facilitates Tuberculosis Invasion in African Buffalo. The American Naturalist 176:613–624.

Ezenwa, V. O., and A. E. Jolles. 2011. From host immunity to pathogen invasion: the effects of helminth coinfection on the dynamics of microparasites. Integrative and comparative biology 51:540–51.

Fenton, A., and S. E. Perkins. 2010. Applying predator-prey theory to modelling immune-mediated, within-host interspecific parasite interactions. Parasitology 137:1027–38.

Fierer, N., S. Ferrenberg, G. E. Flores, A. González, J. Kueneman, T. Legg, R. C. Lynch, D. McDonald, J. R. Mihaljevic, S. P. O'Neill, M. E. Rhodes, S. J. Song, and W. a. Walters. 2012. From Animalcules to an Ecosystem: Application of Ecological Concepts to the Human Microbiome. Annual Review of Ecology, Evolution, and Systematics 43:137–155.

Firth, D. 1993. Bias Reduction of Maximum Likelihood Estimates. Biometrika 80:27–38.

Fischer, J., and D. Lindenmayer. 2002. Treating the nestedness temperature calculator as a "black box" can lead to false conclusions. Oikos:193–199.

Fitzpatrick, M., N. Gotelli, and A. Ellison. 2013. MaxEnt vs. MaxLike: Empirical comparisons with ant species distributions. Ecosphere 4.

Forson, D. D., and A. Storfer. 2006. Atrazine increases ranavirus susceptibility in the tiger salamander, Ambystoma tigrinum. Ecological applications : a publication of the Ecological Society of America 16:2325–32.

Frost, S. D., M. J. Dumaurier, S. Wain-Hobson, and a J. Brown. 2001. Genetic drift and within-host metapopulation dynamics of HIV-1 infection. PNAS 98:6975–80.

G. R. Hess, S. E. Randolph, P. Arneberg, C. Chemini, C. Furlanello, J. Harwood, M. G. Roberts, and J. Swinton. 2002. Spatial aspects of disease dynamics. Page 240 *in* P. J. Hudson, A. Rizzoli, B. T. Grenfell, H. Heesterbeek, and A. P. Dobson, editors. The Ecology of Wildlife Diseases. Oxford University Press.

Gelman, A. 1996. Inference and monitoring convergence. Pages 131–143 Markov Chain Monte Carlo in Practice.

Gelman, A., J. Carlin, H. Stern, and D. Dunson. 2013. Bayesian data analysis. Page 696. Second edition. Chapman and Hall.

Gilpin, M., and J. Diamond. 1982. Factors contributing to non-randomness in species co-occurrences on islands. Oecologia:75–84.

Gleason, H. 1926. The individualistic concept of the plant association. Bulletin of the Torrey Botanical Club 53:7–26.

Gosner, K. L. 1960. A Simplified Table for Staging Anuran Embryos Larvae with Notes on Identification. Herpetologica 16:183–190.

Gotelli, N. J. 2000. Null Model Analysis of Species Co-Occurrence Patterns. Ecology 81:2606.

Gotelli, N. J., and G. R. Graves. 1996. Null Models in Ecology. Smithsonian Institution Press, Washington, D.C.

Graham, A. L. 2008. Ecological rules governing helminth-microparasite coinfection. Proceedings of the National Academy of Sciences of the United States of America 105:566–70.

Granoff, A., P. Came, and K. Rafferty. 1965. The isolation and properties of viruses from Rana pipiens: their possible relationship to the renal adenocarcinoma of the leopard frog. … of the New York Academy of … 5176.

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

Gray, M., D. Miller, and J. Hoverman. 2009a. First Report of Ranavirus Infecting Lungless Salamanders. Herpetological Review 30:316–319.

Gray, M., D. Miller, and J. Hoverman. 2009b. Ecology and pathology of amphibian ranaviruses. Diseases of Aquatic Organisms 87:243–266.

Greer, A. L., J. L. Brunner, and J. P. Collins. 2009. Spatial and temporal patterns of Ambystoma tigrinum virus (ATV) prevalence in tiger salamanders Ambystoma tigrinum nebulosum. Diseases of aquatic organisms 85:1–6.

Greischar, M. a, and B. Koskella. 2007. A synthesis of experimental work on parasite local adaptation. Ecology letters 10:418–34.

Grenfell, B., and J. Harwood. 1997. (Meta)population dynamics of infectious diseases. Trends in Ecology & Evolution 12:395–399.

Griffiths, E. C., A. B. Pedersen, A. Fenton, and O. L. Petchey. 2014. Analysis of a summary network of co-infection in humans reveals that parasites interact most via shared resources Analysis of a summary network of co-infection in humans reveals that parasites interact most via shared resources.

Guo, Q., T. Shaffer, and T. Buhl. 2006. Community maturity, species saturation and the variant diversity-productivity relationships in grasslands. Ecology letters 9:1284–92.

Hechinger, R. F., and K. D. Lafferty. 2005. Host diversity begets parasite diversity: bird final hosts and trematodes in snail intermediate hosts. Proceedings. Biological sciences / The Royal Society 272:1059–66.

Henriques-Silva, R., Z. Lindo, and P. R. Peres-Neto. 2013. A community of metacommunities: exploring patterns in species distributions across large geographical areas. Ecology 94:627–639.

Hess, G. R. 1994. Conservation Corridors and Contagious Disease: A Cautionary Note. Conservation Biology 8:256–262.

Hoeksema, J. D., and S. E. Forde. 2008. A meta-analysis of factors affecting local adaptation between interacting species. The American naturalist 171:275–90.

Holyoak, M., M. A. Leibold, and R. D. Holt. 2005. Metacommunities: spatial dynamics and ecological communities. Page 513. University of Chicago Press.

Hooten, M. B., and N. T. Hobbs. 2014. A Guide to Bayesian Model Selection for Ecologists.

Hoverman, J., M. Gray, and D. Miller. 2010. Anuran susceptibilities to ranaviruses: role of species identity, exposure route, and a novel virus isolate. Diseases of Aquatic Organisms 89:97–107.

Hoverman, J. T., M. J. Gray, N. a Haislip, and D. L. Miller. 2011. Phylogeny, life history, and ecology contribute to differences in amphibian susceptibility to ranaviruses. EcoHealth 8:301–19.

Hoverman, J. T., J. R. Mihaljevic, K. L. D. Richgels, J. L. Kerby, and P. T. J. Johnson. 2012. Widespread co-occurrence of virulent pathogens within California amphibian communities. EcoHealth 9:288–92.

Huang, Z. Y. X., W. F. de Boer, F. van Langevelde, V. Olson, T. M. Blackburn, and H. H. T. Prins. 2013. Species' life-history traits explain interspecific variation in reservoir competence: a possible mechanism underlying the dilution effect. PloS one 8:e54341.

Hudson, P. J., A. Rizzoli, B. T. Grenfell, H. Heesterbeek, and A. P. Dobson (Eds.). 2002. The Ecology of Wildlife Diseases. Page 197. Oxford University Press.

Huston, M. a. 1997. Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. Oecologia 110:449–460.

Iknayan, K. J., M. W. Tingley, B. J. Furnas, and S. R. Beissinger. 2013. Detecting diversity: emerging methods to estimate species diversity. Trends in ecology & evolution:1–10.

Jackson, M., M. Turner, S. Pearson, and A. Ives. 2012. Seeing the forest and the trees: multilevel models reveal both species and community patterns. Ecosphere 3:1–16.

Jancovich, J. K., M. Bremont, J. W. Touchman, and B. L. Jacobs. 2010. Evidence for multiple recent host species shifts among the Ranaviruses (family Iridoviridae). Journal of virology 84:2636–47.

Jancovich, J. K., E. W. Davids, a Seiler, B. L. Jacobs, and J. P. Collins. 2001. Transmission of the Ambystoma tigrinum virus to alternative hosts. Diseases of aquatic organisms 46:159–63.

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

Johnson, P., D. L. Preston, J. T. Hoverman, and B. LaFonte. 2013a. Host and parasite diversity jointly control disease risk in complex communities. Proceedings of the National Academy of Sciences 110:16916–16921.

Johnson, P., J. Rohr, J. Hoverman, E. Kellermanns, J. Bowerman, and K. Lunde. 2012. Living fast and dying of infection: host life history drives interspecific variation in infection and disease risk. Ecology Letters 15:235–42.

Johnson, P., D. Sutherland, J. Kinsella, and K. Lunde. 2004. Review of the trematode genus Ribeiroia (Psilostomidae): ecology, life history and pathogenesis with special emphasis on the amphibian malformation problem. Advances in parasitology.

Johnson, P. T. J., R. B. Hartson, D. J. Larson, and D. R. Sutherland. 2008. Diversity and disease: community structure drives parasite transmission and host fitness. Ecology letters 11:1017–26.

Johnson, P. T. J., and J. T. Hoverman. 2012. Parasite diversity and coinfection determine pathogen infection success and host fitness. Proceedings of the National Academy of Sciences of the United States of America 109:9006–9011.

Johnson, P. T. J., D. L. Preston, J. T. Hoverman, and K. L. D. Richgels. 2013b. Biodiversity decreases disease through predictable changes in host community competence. Nature 494:230–3.

Jolles, A. E., V. O. Ezenwa, R. S. Etienne, W. C. Turner, and H. Olff. 2008. Interactions between macroparasites and microparasites drive infection patterns in free-ranging African buffalo. Ecology 89:2239–50.

Joseph, M. B., J. R. Mihaljevic, A. L. Arellano, J. G. Kueneman, D. L. Preston, P. C. Cross, and P. T. J. Johnson. 2013a. Taming wildlife disease: bridging the gap between science and management. Journal of Applied Ecology 50:702–712.

Joseph, M. B., J. R. Mihaljevic, S. a Orlofske, and S. H. Paull. 2013b. Does life history mediate changing disease risk when communities disassemble? Ecology letters 16:1405–12.

Kattenbelt, J., A. Hyatt, and A. Gould. 2000. Recovery of ranavirus dsDNA from formalin-fixed archival material. Diseases of aquatic organisms 39:151–154.

Keele, L. J. 2008. Semiparametric Regression for the Social Sciences. Page 230. John Wiley & Sons.

Keesing, F., L. K. Belden, P. Daszak, A. Dobson, C. D. Harvell, R. D. Holt, P. Hudson, A. Jolles, K. E. Jones, C. E. Mitchell, S. S. Myers, T. Bogich, and R. S. Ostfeld. 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. Nature 468:647–652.

Keesing, F., R. D. Holt, and R. S. Ostfeld. 2006. Effects of species diversity on disease risk. Ecology letters 9:485–98.

Kilpatrick, a M., P. Daszak, M. J. Jones, P. P. Marra, and L. D. Kramer. 2006. Host heterogeneity dominates West Nile virus transmission. Proceedings. Biological sciences / The Royal Society 273:2327–33.

Knowles, S., A. Fenton, O. L. Petchey, T. Jones, R. Barber, and A. B. Pedersen. 2013. Stability of within-host−parasite communities in a wild mammal system. Proceedings of the Royal Society B-Biological Sciences 280.

Krasnov, B. R., D. Mouillot, G. I. Shenbrot, I. S. Khokhlova, M. V Vinarski, N. P. Korallo-Vinarskaya, and R. Poulin. 2010. Similarity in ectoparasite faunas of Palaearctic rodents as a function of host phylogenetic, geographic or environmental distances: which matters the most? International journal for parasitology 40:807–17.

Krasnov, B. R., and M. V Vinarski. 2008. Searching for general patterns in parasite ecology : host identity versus environmental influence on gamasid mite assemblages in small mammals. Parasitology:229–242.

Kuris, A., and A. Blaustein. 1980. Hosts as islands. The American Naturalist 116:570–586.

De la Sancha, N. U., C. L. Higgins, S. J. Presley, and R. E. Strauss. 2014. Metacommunity structure in a highly fragmented forest: has deforestation in the Atlantic Forest altered historic biogeographic patterns? Diversity and Distributions.

Lafferty, K. D. 2010. Interacting Parasites. Science 330:187–188.

Lafferty, K. D., and C. L. Wood. 2013. It's a myth that protection against disease is a strong and general service of biodiversity conservation: Response to Ostfeld and Keesing. Trends in ecology & evolution 28:503–4.

Lahoz-Monfort, J. J., G. Guillera-Arroita, and B. a. Wintle. 2014. Imperfect detection impacts the performance of species distribution models. Global Ecology and Biogeography 23:504–515.

Lehman, C., and D. Tilman. 2000. Biodiversity, stability, and productivity in competitive communities. The American Naturalist 156:534–552.

Leibold, M. A. 2011. The Metacommunity Concept and its Theoretical Underpinnings. Pages 163–184 *in* S. M. Scheiner and M. R. Willig, editors. The Theory of Ecology. University of Chicago Press.

Leibold, M. a., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, and a. Gonzalez. 2004a. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.

Leibold, M. A., and G. M. Mikkelson. 2002. Coherence, species turnover, and boundary clumping: elements of meta-community structure. Oikos 97:237–250.

Leibold, M., M. Holyoak, N. Mouquet, P. Amarasekare, J. Chase, M. Hoopes, R. Holt, J. Shurin, R. Law, D. Tilman, and others. 2004b. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.

Leo, G. De, and A. Dobson. 1996. Allometry and simple epidemic models for microparasites. Nature 379:720–722.

Lesbarrères, D., a Balseiro, J. Brunner, V. G. Chinchar, a Duffus, J. Kerby, D. L. Miller, J. Robert, D. M. Schock, T. Waltzek, and M. J. Gray. 2012. Ranavirus: past, present and future. Biology letters 8:481–3.

Lindström, E. S., and S. Langenheder. 2011. Local and regional factors influencing bacterial community assembly. Environmental Microbiology Reports.

Loeuille, N., and M. a Leibold. 2008. Evolution in metacommunities: on the relative importance of species sorting and monopolization in structuring communities. The American naturalist 171:788–99.

LoGiudice, K., S. T. K. Duerr, M. J. Newhouse, K. a Schmidt, M. E. Killilea, and R. S. Ostfeld. 2008. Impact of host community composition on Lyme disease risk. Ecology 89:2841–9.

LoGiudice, K., R. S. Ostfeld, K. a Schmidt, and F. Keesing. 2003. The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. Proceedings of the National Academy of Sciences of the United States of America 100:567–71.

Logue, J. B., N. Mouquet, H. Peter, and H. Hillebrand. 2011. Empirical approaches to metacommunities: a review and comparison with theory. Trends in ecology & evolution 26.

López-González, C., S. J. Presley, A. Lozano, R. D. Stevens, and C. L. Higgins. 2012. Metacommunity analysis of Mexican bats: environmentally mediated structure in an area of high geographic and environmental complexity. Journal of Biogeography 39:177–192.

López-Villavicencio, M., F. Courjol, A. K. Gibson, M. E. Hood, O. Jonot, J. a Shykoff, and T. Giraud. 2011. Competition, cooperation among kin, and virulence in multiple infections. Evolution 65:1357–66.

MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Page 203. Princeton University Press.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.

McCallum, H. 2012. Disease and the dynamics of extinction. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 367:2828–39.

McCallum, H., and A. Dobson. 2002. Disease, habitat fragmentation and conservation. Proceedings. Biological sciences / The Royal Societyof the Royal 269:2041–2049.

De Meester, L. 2011. A metacommunity perspective on the phylo-and biogeography of small organisms. Pages 324–333 Biogeography of Microscopic Organisms: Is Everything Small Everywhere? Cambridge Univ Pr.

Meynard, C. N., I. Boulangeat, L. Garraud, N. Mouquet, and W. Thuiller. 2013. Disentangling the drivers of metacommunity structure across spatial scales:1–12.

Mideo, N. 2009. Parasite adaptations to within-host competition. Trends in parasitology 25:261–8.

Mideo, N., S. Alizon, and T. Day. 2008. Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. Trends in Ecology & Evolution 23:511–7.

Mihaljevic, J. R. 2012. Linking metacommunity theory and symbiont evolutionary ecology. Trends in ecology & evolution.

Miller, D., and S. Rajeev. 2007. Frog Virus 3 Infection, Cultured American Bullfrogs. Emerging infectious … 13:2006–2007.

Mitchell, C., D. Tilman, and J. Groth. 2002. Effects of grassland plant species diversity, abundance, and composition on foliar fungal disease. Ecology 83:1713–1726.

Morand, S., and R. Poulin. 1998. Density , body mass and parasite species richness of terrestrial mammals.

Van Nood, E., A. Vrieze, M. Nieuwdorp, S. Fuentes, E. G. Zoetendal, W. M. de Vos, C. E. Visser, E. J. Kuijper, J. F. W. M. Bartelsman, J. G. P. Tijssen, P. Speelman, M. G. W. Dijkgraaf, and J. J. Keller. 2013. Duodenal infusion of donor feces for recurrent Clostridium difficile. The New England journal of medicine 368:407–15.

Nunn, C. L., C. Brezine, A. E. Jolles, and V. O. Ezenwa. 2014. Interactions between micro- and macroparasites predict microparasite species richness across primates. The American naturalist 183:494–505.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2013. vegan: Community Ecology Package.

Osnas, E. E., and A. P. Dobson. 2011. Evolution of virulence in heterogeneous host communities under multiple trade-offs. Evolution.

Ostfeld, R. S., and F. Keesing. 2000. The function of biodiversity in the ecology of vector-borne zoonotic diseases. Canadian Journal of Zoology 78:2061–2078.

Ostfeld, R. S., and F. Keesing. 2012. Effects of Host Diversity on Infectious Disease. Annual Review of Ecology, Evolution, and Systematics 43:157–182.

Ostfeld, R. S., F. Keesing, and V. T. Eviner (Eds.). 2008. Infectious Disease Ecology: Effects of Ecosystems on Disease and of Disease on Ecosystems. Page 506. Princeton University Press.

Pedersen, A. B., S. Altizer, M. Poss, A. a Cunningham, and C. L. Nunn. 2005. Patterns of host specificity and transmission among parasites of wild primates. International journal for parasitology 35:647–57.

Pedersen, A. B., and A. Fenton. 2006. Emphasizing the ecology in parasite community ecology. Trends in Ecology & Evolution 22:133–139.

Petney, T. N., and R. H. Andrews. 1998. Multiparasite communities in animals and humans: frequency, structure and pathogenic significance. International journal for parasitology 28:377–93.

Picco, A. M., J. L. Brunner, and J. P. Collins. 2007. Susceptibility of the endangered California tiger salamander, Ambystoma californiense, to ranavirus infection. Journal of wildlife diseases 43:286–90.

Pillai, P., A. Gonzalez, and M. Loreau. 2011. Metacommunity theory explains the emergence of food web complexity. Proceedings of the National Academy of Sciences of the United States of America 108:19293–19298.

Poulin, R. 2004. Parasites and the neutral theory of biodiversity. Ecography 27:119–123.

Poulin, R. 2007. Are there general laws in parasite ecology? Parasitology 134:763–76.

Poulin, R., B. R. Krasnov, D. Mouillot, and D. W. Thieltges. 2011. The comparative ecology and biogeography of parasites. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 366:2379–90.

Presley, S. J., C. L. Higgins, and M. R. Willig. 2010. A comprehensive framework for the evaluation of metacommunity structure. Oikos 119:908–917.

Presley, S. J., and M. R. Willig. 2010. Bat metacommunity structure on Caribbean islands and the role of endemics. Global Ecology and Biogeography 19:185–199.

Preston, F. 1948. The Commonness, And Rarity, of Species. Ecology 29:254–283.

Prudhoe, S., and R. Bray. 1982. Platyhelminth parasites of the Amphibia.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Randolph, S. E., and a D. M. Dobson. 2012. Pangloss revisited: a critique of the dilution effect and the biodiversity-buffers-disease paradigm. Parasitology 139:847–63.

Read, A. F., and L. H. Taylor. 2001. The Ecology of Genetically Diverse Infections. Science 292:1099–1102.

Reperant, L. 2010. Applying the Theory of Island Biogeography to Emerging Pathogens: Toward Predicting the Sources of Future Emerging Zoonotic and Vector-Borne Diseases. Vector-Borne and Zoonotic Diseases 10:105–110.

Richgels, K. L. D., J. T. Hoverman, and P. T. J. Johnson. 2013. Evaluating the role of regional and local processes in structuring a larval trematode metacommunity of Helisoma trivolvis. Ecography 36:854–863.

Ricklefs, R. E. 2008. Disintegration of the ecological community. The American Naturalist 172:741–50.

Ridenhour, B. J., and A. T. Storfer. 2008. Geographically variable selection in Ambystoma tigrinum virus (Iridoviridae) throughout the western USA. Journal of evolutionary biology 21:1151–9.

Rigaud, T., M.-J. Perrot-Minnot, and M. J. F. Brown. 2010. Parasite and host assemblages: embracing the reality will improve our knowledge of parasite transmission and virulence. Proceedings. Biological sciences / The Royal Society 277:3693–3702.

Roberts, L., J. Janovy, and G. Schmidt. 1996. Foundations of parasitology.

Roche, B., A. P. Dobson, J.-F. Guégan, and P. Rohani. 2012. Linking community and disease ecology: the impact of biodiversity on pathogen transmission. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 367:2807–13.

Roche, B., P. Rohani, A. P. Dobson, and J.-F. Guégan. 2013. The Impact of Community Organization on Vector-Borne Pathogens. The American Naturalist 181:1–11.

De Roode, J. C., R. Pansini, S. J. Cheesman, M. E. H. Helinski, S. Huijben, A. R. Wargo, A. S. Bell, B. H. K. Chan, D. Walliker, and A. F. Read. 2005. Virulence and competitive ability in genetically diverse malaria infections. Proceedings of the National Academy of Sciences of the United States of America 102:7624–8.

Rosenzweig, M. 1995. Species diversity in space and time.

Royle, J., R. Chandler, C. Yackulic, and J. D. Nichols. 2012. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. Methods in Ecology and Evolution 3:545–554.

Royle, J., and R. Dorazio. 2008. Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations, and Communities. Page 464.

Rudolf, V. H. W., and J. Antonovics. 2005. Species coexistence and pathogens with frequency-dependent transmission. The American naturalist 166:112–8.

Schock, D. M., T. K. Bollinger, V. Gregory Chinchar, J. K. Jancovich, and J. P. Collins. 2008. Experimental Evidence that Amphibian Ranaviruses Are Multi-Host Pathogens. Copeia 2008:133–143.

Sloan, W. T., M. Lunn, S. Woodcock, I. M. Head, S. Nee, and T. P. Curtis. 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. Environmental microbiology 8:732–40.

Smith, K. F., D. F. Sax, and K. D. Lafferty. 2006. Evidence for the role of infectious disease in species extinction and endangerment. Conservation biology 20:1349–57.

Smyth, J., and M. Smyth. 1980. Frog as Host-Parasite Systems 1: An Introduction to Parasitology through the Parasites of Rana temporaria, R. esculenta and R. pipiens. The Macmillan Press LTD, London.

Svensson-Coelho, M., and R. E. Ricklefs. 2011. Host phylogeography and beta diversity in avian haemosporidian (Plasmodiidae) assemblages of the Lesser Antilles. The Journal of animal ecology 80:938–946.

Telfer, S., X. Lambin, R. Birtles, P. Beldomenico, S. Burthe, S. Paterson, and M. Begon. 2010. Species Interactions in a Parasite Community Drive Infection Risk in a Wildlife Population. Science 330:243.

Thompson, R. C. a, a J. Lymbery, and a Smith. 2010. Parasites, emerging disease and wildlife conservation. International journal for parasitology 40:1163–70.

Thrall, P. H., and J. J. Burdon. 1997. Host-pathogen dynamics in a metapopulation context: the ecological and evolutionary consequences of being spatial. Journal of Ecology 85:743–753.

Tilman, D., P. B. Reich, J. Knops, D. Wedin, T. Mielke, and C. Lehman. 2001. Diversity and productivity in a long-term grassland experiment. Science (New York, N.Y.) 294:843–5.

Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature 379:718–720.

Tompkins, D. M., A. M. Dunn, M. J. Smith, and S. Telfer. 2010. Wildlife diseases: from individuals to ecosystems. The Journal of animal ecology 4:1–20.

Ulrich, W., and N. J. Gotelli. 2013. Pattern detection in null model analysis. Oikos 122:2–18.

Urban, M. C. 2011. The evolution of species interactions across natural landscapes. Ecology Letters 14:723–732.

Urban, M. C., M. A. Leibold, P. Amarasekare, L. De Meester, R. Gomulkiewicz, M. E. Hochberg, C. a Klausmeier, N. Loeuille, C. de Mazancourt, J. Norberg, J. H. Pantel, S. Y. Strauss, M. Vellend, and M. J. Wade. 2008. The evolutionary ecology of metacommunities. Trends in Ecology & Evolution 23:311–7.

Urban, M. C., and L. De Meester. 2009. Community monopolization: local adaptation enhances priority effects in an evolving metacommunity. Proceedings. Biological sciences / The Royal Society 276:4129–38.

Urban, M. C., L. De Meester, M. Vellend, R. Stoks, and J. Vanoverbeke. 2011. A crucial step toward realism: responses to climate change from an evolving metacommunity perspective. Evolutionary Applications.

Urban, M. C., and D. K. Skelly. 2006. EVOLVING METACOMMUNITIES: TOWARD AN EVOLUTIONARY PERSPECTIVE ON METACOMMUNITIES. Ecology 87:1616–1626.

Watanabe, S. 2013. A widely applicable Bayesian information criterion. The Journal of Machine Learning Research 14:867–897.

Wood, C. L., and K. D. Lafferty. 2012. Biodiversity and disease: a synthesis of ecological perspectives on Lyme disease transmission. Trends in ecology & evolution:1–9.

Wood, S. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B (… 73:3–36.

Youngster, I., J. Sauk, C. Pindar, R. G. Wilson, J. L. Kaplan, M. B. Smith, E. J. Alm, D. Gevers, G. H. Russell, and E. L. Hohmann. 2014. Fecal microbiota transplant for relapsing Clostridium difficile infection using a frozen inoculum from unrelated donors: a randomized, open-label, controlled pilot study. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 58:1515–22.

Zwetyenga, J., C. Rogier, a Tall, D. Fontenille, G. Snounou, J. F. Trape, and O. Mercereau-Puijalon. 1998. No influence of age on infection complexity and allelic distribution in Plasmodium falciparum infections in Ndiop, a Senegalese village with seasonal, mesoendemic malaria. The American journal of tropical medicine and hygiene 59:726–35.

APPENDIX

Appendix: Chapter 2, Generating the "saturating" method and null model

A.1 The "saturating" method

Communities were assembled by the "saturating" method in a way that combined additive and compensatory species additions. Our goal was to simulate nearly additive species additions at low values of richness and then transition to compensatory additions at higher values of richness. To achieve this, species' abundances were adjusted by a scaling factor, termed K_s, so that the total community density, K_T, varied with species richness according to a saturating function, $K_{com}(R)$, where R refers to host species richness. This ensured that species abundances were adjusted in proportion to their equilibrial abundances in the global pool. The two saturating functions used in the main body text were:

$$K_{com} = 500 - \frac{3100}{(R+5)} \quad \text{(A1)},$$

and:

$$K_{com} = \frac{500}{1+50e^{-0.15(R+10)}} \quad \text{(A2)}.$$

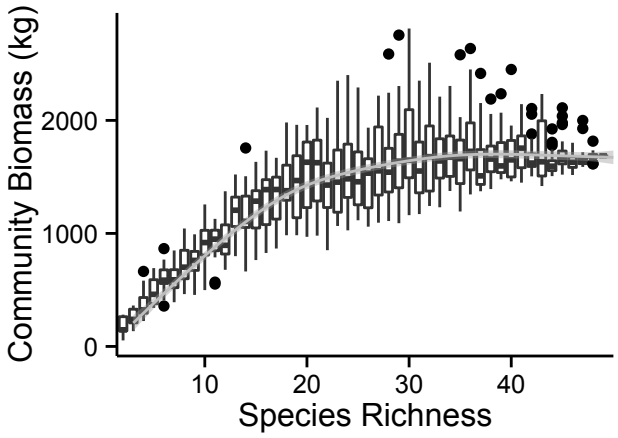
Equation (A1) is a typical asymptotic function, and is similar in form to an empirically derived relationship between plant percent cover and species richness (Tilman et al. 1996). Equation (A2) represents a logistic growth curve. The scaling factor was then calculated as:

$$K_s = {K_{com}}/{K_T} \quad \text{(S3)}.$$

K_s was multiplied by each species' equilibrial abundance in the assembled community in order to calculate adjusted abundances. However, in order to introduce more variation and additive increases in abundance, if $K_T < K_{com}$, equilibrial abundances were not adjusted. Therefore, at low host richness, species additions were mostly additive but gradually transitioned to completely compensatory additions at high richness. This method of saturating communities also

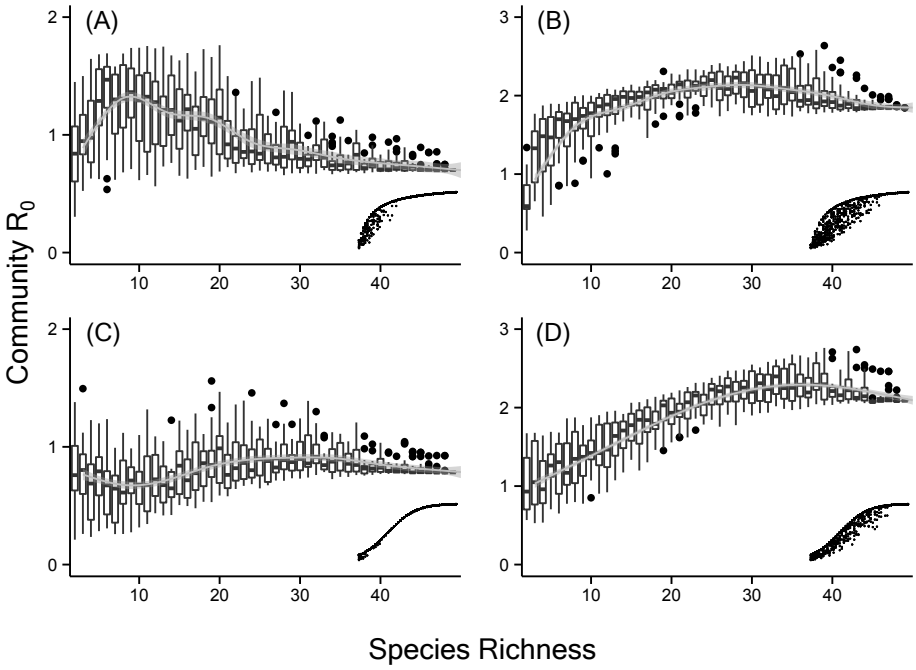
corresponds to patterns found in the ecosystem function literature that show a saturating relationship between total community biomass and species richness (Lehman and Tilman 1997; Tilman et al. 2001; Guo et al. 2006; Figure A.1).

Figure A.1. An example of the relationship between total community biomass and species richness produced by the "saturating" method with 1000 simulated communities. Boxplots summarize the data for each value of richness. This example corresponds to the case where the saturating relationship is as in equation (A1). A LOESS smoother and 95% confidence bands were added for visual interpretation of the average trend.



Whether the "saturating" method led to strong or weak non-monotonic relationships between community R_0 and species richness was somewhat sensitive to the exact formulation of the saturating function, but that the non-monotonic relationship was general. Specifically, the maximal host community abundance (e.g. ~500 in equations S1 and S2) mediated the severity of the "hump" shaped relationship. Higher maximal host community abundance, which represents weaker compensatory interactions, led to a less pronounced hump, while lower maximal host abundance showed the opposite trend (Figure A.2).

Figure A.2. Community R_0 versus species richness with various saturating functions, all assuming density-dependent transmission. *A*-*B*, Variations of equation 1 that alter the maximal host community abundance to ~300 (*A*) and ~800 (*B*). *C*-*D*, Variations of equation 2 that alter the maximal community abundance to ~300 (*C*) and ~800 (*D*). Inset figures represent the underlying community abundance-richness relationships. A LOESS smoothing line and 95% confidence bands are added only for visual interpretation. Exact equations for K_{com} are as follows: *A*, $K_{com} = 300 - 1800/(R+5)$; *B*, $K_{com} = 800 - 5300/(R+5)$; *C*, $K_{com} = \frac{300}{1+50e^{-0.15(R+10)}}$; *D*, $K_{com} = \frac{800}{1+50e^{-0.15(R+10)}}$.



A.2 Null model results

For this model, we randomized all life-history traits to eliminate associations with intraspecific R_0, R_{0i}. We then derived intraspecific transmission rates β_{ii} to match R_{0i} for each species and used this community to generate our simulation scenarios. We find that the assumption that the most abundant species is also the most competent host does not affect our qualitative results (Figure A.3).

Figure A.3. Selected results of the 'null' model in which there are random associations between host competence, abundance, and other life-history traits. *A*, "additive" abundance-richness relationship; *B*, "fixed" abundance-richness relationship; and *C*, "saturating" abundance-richness. All panels were simulated with density-dependent transmission.