Characteristics of the Emergent Disease *Batrachochytrium dendrobatidis* in the *Rana muscosa* and *Rana sierrae* Species Complex

by

Tate Scott Tunstall

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Cheryl Briggs, Chair
Professor Craig Moritz, Co-chair
Professor Montgomery Slatkin
Professor Wayne Getz

Fall 2012

Characteristics of the Emergent Disease *Batrachochytrium dendrobatidis* in the *Rana muscosa* and *Rana sierrae* Species Complex

Abstract

Characteristics of the Emergent Disease *Batrachochytrium dendrobatidis* in the *Rana muscosa* and *Rana sierrae* Species Complex

by

Tate Scott Tunstall

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Cheryl Briggs, Chair

Professor Craig Moritz, Co-chair

The fungal pathogen *Batrachochytrium dendrobatidis*, or Bd, has been a major driver of amphibian extinctions world wide. This dissertation investigates the effects of Bd on the mountain yellow legged frogs, *Rana muscosa* and *Rana sierrae*. These two species occur in the Sierra Nevada in California, and have under gone dramatic declines in part due to the invasion of Bd. Even before the invasion of Bd, populations of *Rana muscosa/sierrae* have faced habitat loss and fragmentation from introduced predators and habitat fragmentation. The fist chapter of this dissertation investigates how genetic diversity effects R_0, the threshold for invasion of a pathogen. I model the case where a naive population faces multiple disease, as well as a single population faced with invasion of a single disease. I also model the effect of overdominance and genetic diversity on R_0. In my second chapter, I use both models and simulations to investigate how genetic diversity effects the final size of an epidemic, and whether or not populations with higher genetic diversity maintain that diversity after an epidemic. In my third chapter, I use microsatellite markers to measure genetic diversity in serveral populations of *Rana muscosa* and *Rana sierrae*, and test whether or not these populations have experienced historic bottlenecks. In my final chapter, I present data on a series of experiments on how variation in the external source of Bd, or zoospore pool, affects the growth rate of Bd on its amphibian host, as well as host mortality.

To Maya

Never could have done it without you

Contents

List of Figures

List of Tables

Acknowledgments

I want to thank all of my committe members: Wayne Getz, Monty Slatkin, and Craig Moritz. Your comments and advice have been invaluable. I would also like to thank all the faculty and UCB and UCSB that I have interacted with over the years. I've truely been lucky to get to experience two of the finest educational institutions in the world. I'd especially like to thank my advisor Cherie Briggs, for taking me as on as a student, and being a better advisor than I could ever hope for. I can't leave out the rest of the Briggs lab, for all the good times in Berkeley and Santa Barbara. My dear friends and fellow graduate students Sean Schoville, Sean Rovito, and Tom Devitt, without whose advice I would've never made it through school. Special thanks to my friend and mentor Vance. My mother, who has supported my always, my brother Trent, who's always been there for me, and of course my wife, Maya.

Chapter 1

The effect of genetic heterogeneity on R_0

1.1 Introduction

When a disease invades a naive population, there can be considerable variation in individual susceptibility, tolerance, and transmission. There may be some level of natural resistance in the population, with some individuals being completely or partially resistant to infection. "Superspreaders" are infected individuals who are responsible for a disproportionately large amount of the transmission [Lloyd-Smith et al., 2006], such individuals appear to have played a significant role in the recent SARS outbreak [Shen et al., 2004]. Much of this variation is doubtless due to factors such as heterogeneity in contact networks, behavior, and environment. Models have shown, for example, that disease can spread more rapidly when there is heterogeneity in contact between individuals [Diekmann et al., 1990, Lloyd-Smith et al., 2006]. This paper will instead focus on how *genetic* heterogeneity may contribute to the severity of disease outbreaks. Specifically, we focus on how heterogeneity tied to genetic variation affects both R_0, a metric used in many epidemiological modles describing the average number of secondary infections caused by a single infected individual in a completely susceptible population. Many epidemiological models describe their results in terms of R_0: when $R_0 < 1$ a disease is unlikely to become an epidemic, whereas when $R_0 > 1$ an epidemic is probable.

Recently, there have been several studies examining how genetic heterogeneity contributes to an epidemic [Springbett et al., 2003, Doeschl-Wilson et al., 2011], and recent advances in sequencing techniques have shed light on how individual genes affect disease parameters [Lohmueller et al., 2003, Hirschhorn and Daly, 2005, Daetwyler et al., 2008]. There is still much to be learned about how variation at the genetic level translates to heterogeneity at the population level, and thus previous attempts to model genetic heterogeneity have taken a variety of approaches. Springbett et al. [2003] modeled heterogeneity in susceptibility, and found that genetic heterogeneity has no impact on the expected value of R_0. Doeschl-Wilson

et al. [2011] compared heterogeneous populations with the same average susceptibility and found that disease risk and severity were higher in populations containing larger proportions of susceptible individuals. Lively [2010] modeled a population exposed to multiple pathogens, each with a specific resistance genotype, and found that the average $\bar{R_0}$ is inversely related to the number of host genotypes in the population.

Overdominance, where heterozygote genotypes have higher fitness than homozygotes, has been suggested as a mechanism that maintains genetic diversity in the major histocompatability complex , (or MHC), genes that are an important component of the immune system [Wegner et al., 2004, Slade and McCallum, 1992]. Overdominance is believed to be important in resistance to malaria [Modiano et al 2001] and hepatitis C [Hraber et al., 2007]. Although it is difficult to distinguish frequency dependent selection from overdominance [Takahata and Nei, 1990], genetic polymorphism resulting from overdominance should be stable, while polymorphism from frequency dependent selection will be dynamic [Slade and McCallum, 1992]. The overdominance hypothesis for maintenance of MHC diversity states that individuals with more MHC alleles (heterozygotes) should have higher fitness, as they are capable of fighting off a greater number of diseases [Kamath and Getz, 2011].

Here we investigate the impact of heterogeneity in susceptibility due to genetic variation on R_0. Specifically, we will investigate the possibly countervailing effects of an increase in population size along with a corresponding increase in genetic diversity. In diseases with density dependent transmission, increasing population size in a fixed area (i.e. population density) will increase R_0 through the increased number of contacts between individuals. Genetic diversity will also increase with population size, both in terms of the number of unique alleles in the population and the number of individuals heterozygous at a given locus. Increasing population size, and diversity of alleles in the population, will increase the probability that a resistant genotype is present in a population. While having more resistant genotypes in a population will most certainly reduce the probability and severity of an epidemic, this is not an advantage of heterogeneity *per se*, as homogeneous populations consisting solely of resistant genotypes will be have a lower R_0 than heterogeneous populations with a few resistance alleles. The purpose of this study is to determine how genetic heterogeneity itself affects a population, specifically how an increase in heterogeneity due to an increase in population size affects the basic reproductive number R_0.

1.2 Methods

Disease Model

The model used in this study was a modification of the basic SIR (Susceptible, Infected, Recovered/Removed) compartmental model frequently used to describe infectious disease dynamics [Anderson and May, 1991]:

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= -\beta SI \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \beta SI - \gamma I \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma I\end{aligned}$$

The above model describes an epidemic of a directly transmitted disease through a population in which S, I and R are the numbers of susceptible, infected and resistant individuals, respectively. The model assumes that the total population size $N = S + I + R$, where β is the transmission rate between infected and susceptibles, and γ is rate of recovery into the resistant/removed class. The model describes only short term dynamics, and does not include births or deaths. The basic reproductive ratio R_0, using the above model, is $\beta N/\gamma$. In our model we assume that the susceptibile class can be divided into k subgroups, each labeled S_i. We break down the transmission process β into infectivity ϕ, susceptibility σ, and contact rate c, and consider that each of these parameters could potentially vary by genotype:

$$\begin{aligned}\frac{\mathrm{d}S_i}{\mathrm{d}t} &= -\sum_j \sigma_i \phi_j c_{ij} S_i I_j \\ \frac{\mathrm{d}I_i}{\mathrm{d}t} &= \sum_j \sigma_i \phi_j c_{ij} S_i I_j - \gamma_i I_i \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \sum_i \gamma_i I_i\end{aligned}$$

We define susceptibility as the probability of becoming infected given an infectious contact with an infected individual, and infectivity as the probability that a contact is potentially infectious, given a contact with a susceptible. Full contact between groups was assumed with random mixing of all individuals, and we assume separable mixing, ie the absence of correlation between infectivity and susceptibility between individuals (as opposed to within individuals) [Diekmann et al., 1990]. In other words, transmission between a susceptible individual of type i with properties a_i and an infected individual of type j with properties b_j is equal to $a_i b_j$. In a population with separable mixing, R_0 is just the average R_0 over all subpopulations $\sum \beta_i N_i/\gamma_i$. In the model below, we will consider both density dependent and frequency dependent transmission. For the purposes of this paper, we will define resistance to a disease as a change in any of the parameters that affect transmission, such as a decrease in susceptibility σ or an increase in the recovery rate γ, for example.

Genetic Model

A population facing a novel pathogen will likely have little natural resistance. If a pathogen is truly novel, then any variation in resistance to the pathogen will likely be neutral: in a

naive population there is no selection for or against resistance, assuming the cost of resistance is low. We assume mutations between alleles are specified by a k-allele model, where the mutation rate between alleles is specified by a parameter μ, and an allele i mutates into each of the $k-1$ other allelic states with equal probability. This results in a k-allele Wright-Fisher model [Ewens, 2004]. We assume that the parameters are determined by a single locus, and any covariance in parameters is due to effects at the single locus. We assumed that when a population is free of disease, any variation in the population is maintained through neutral evolution, and that at mutation drift equilibrium, the homozygosity, or the fraction of the population that are homozygotes, F, of a population is given by [Ewens, 2004]:

$$F = \frac{k-1+\theta}{k-1+k\theta}$$

where θ, the population mutation rate, is equal to $4N\mu$, where N is the total population size and μ is the mutation rate. We assume populations have not been previously exposed to our pathogen, and there has been no selection for resistance. In other words, the expected frequency of an allele i conferring some degree of resistance is independent of the amount of resistance conferred.

1.3 Results

Resistance to Multiple diseases

In this section, we will consider how diversity affects R_0 in a naive populaiton facing multiple diseases. When considering resistance to multiple diseases, we assume that an individual with one copy of an allele i has the same transmission properties as an individual with two copies, ie an additional copy of an allele confers no additional resistance. Each allele confers resistance to only a single disease, also denoted by i out of z possible diseases. We assume that the contact rate c does not vary ($c_i = c$ for all i), and that the allele can affect the susceptibility σ_i, infectivity ϕ_i, and duration of infection $1/\gamma_i$ of members of the susceptibility class S_i, where parameters marked with hat notation ($\hat{.}$) represent the baseline (disease parameters unaffected by genotype) susceptibility ($\hat{\sigma}_i$), infectivity ($\hat{\phi}_i$), and duration of infection ($1/\hat{\gamma}_i$), of a disease i. In other words, $\hat{S}_i$ refers to the number of susceptible individuals with no resistance exposed to *disease* i, and S_i refers to susceptibles with the resistance *allele* i. For simplicity, in this section we assume all diseases have density dependent transmission, where the transmission rate is equal to βSI.

First, we calculate R_{0i}, which is R_0 for a disease i:

$$\begin{aligned} R_{0i} &= \hat{S}_i \frac{\hat{\sigma}_i c\hat{\phi}_i}{\hat{\gamma}_i} + S_i \frac{\sigma_i c\phi_i}{\gamma_i} \\ &= (N - S_i)\frac{\hat{\sigma}_i c\hat{\phi}_i}{\hat{\gamma}_i} + S_i \frac{\sigma_i c\phi_i}{\gamma_i} \\ &= \hat{R}_{0i} - S_i\Delta_i \end{aligned}$$

where $\hat{R}_{0i}$, is R_0 in a completely susceptible population exposed to disease i, and Δ_i is the difference in transmission multiplied by the duration of infection between an individual with genotype i and individual without resistance exposed to disease i:

$$\Delta_i = \frac{\sigma_i c\phi_i}{\gamma_i} - \frac{\hat{\sigma}_i c\hat{\phi}_i}{\hat{\gamma}_i}$$

Δ_i will be negative when an allele confers a high level or resistance. Now, if we assume z possible diseases and k possible resistance alleles, we calculate the mean value of R_0 over all diseases ($E_D\{R_0\}$), where each disease is given equal weight:

$$\begin{aligned} E_D\{R_0\} &= \frac{1}{z}\sum_{i=1}^{z}(\hat{R}_{0i} - S_i\Delta_i) \\ &= E_D\{\hat{R}_0\} - \frac{1}{z}\sum_{i=1}^{k} S_i\Delta_i \end{aligned} \tag{1.1}$$

where $E_D\{\hat{R}_0\}$ is the average value of R_0 in a population with no resistance alleles exposed to disease. S_i, the susceptible population with allele i, can be broken down into homozygous and heterozygous individuals, therefore we have:

$$\begin{aligned} \frac{1}{z}\sum_{i=1}^{k} S_i\Delta_i &= \frac{1}{z}\sum_{i=1}^{k}(Nf_{o,i} + Nf_{e,i})\Delta_i \\ &= \frac{N}{z}\sum_{i=1}^{k}(f_{o,i} + f_{e,i})\Delta_i \end{aligned} \tag{1.2}$$

Where $f_{o,i}$ and $f_{e,i}$ are the proportions of homozygous and heterozygous individuals with at least one i allele, respectively. Given our assumption of no selection on disease alleles, and that allele mutates into any other state with equal probability, the frequencies of $f_{o,i}$ and $f_{e,i}$ do not depend on the allele i or Δ_i. For each allele i, there is one homozygous genotype and $k-1$ heterozygous genotypes, with each homozygote and heterozygote genotype having

the same expected frequency, respectively. Combining equations (1.1) and (1.2), and then taking the expectation over all possible populations and allele frequencies, $E_{D,f}\{R_0\}$:

$$\begin{aligned} E_{D,f}\{R_0\} &= E_D\{\hat{R}_0\} - \frac{N}{z}(\bar{f}_o + (k-1)\bar{f}_e)\sum_{i=1}^{k}\Delta_i \\ &= E_D\{\hat{R}_0\} - \frac{N}{z}(\bar{f}_o + (k-1)\bar{f}_e)k\bar{\Delta} \end{aligned}$$

where $E_{D,f}\{R_0\}$ is R_0 averaged over all diseases and all allele frequencies. $k\bar{f}_o$ is equal to F, the homozygosity of the population, because there are k possible homozygous genotypes, each with the same expected frequency under our model. Similarly, $(k^2-k)f_e$ is equal to $2H$, where H is the heterozygosity of the population, because there are $(k^2-k)/2$ possible heterozygous genotypes. Given that H is $1-F$, we have:

$$E_{D,f}\{R_0\} = E_D\{\hat{R}_0\} - \frac{N}{z}\bar{\Delta}\,(2-F) \tag{1.3}$$

As the population size N increases, the homozygosity F decreases. As F goes to zero, equation (1.3) becomes:

$$E_{D,f}\{R_0\} = E_D\{\hat{R}_0\} - \frac{2N\bar{\Delta}}{z} \tag{1.4}$$

where z is the number of possible diseases. Increasing N will always increase R_0 and the increase in transmission due to increase in population density will swamp any reduction due to increase in heterozygosity. When the heterozygotes are resistant to multiple diseases, and homozygotes only resitant to only one disease, with no additional benefit from a second allele, the increase in R_0 will be nonlinear as F goes to zero. The more diseases an organism can face, the less effective this reduction will be, providing no benefit as z goes to infinity.

Comparing the above results to a haploid population, where the heterozygote frequency is zero, and $\bar{f}_o$ is equal to $1/k$, equation (1.4) reduces to:

$$E_{D,f}\{R_0\} = E_D\{\hat{R}_0\} - \frac{N\bar{\Delta}}{z} \tag{1.5}$$

In the above equation, $E_{D,f}\{R_0\}$ is equal to the expected value of R_0 in a completely susceptible population with no genetic resistance, reduced by the average reduction in susceptibility due to genotype, and increases linearly with N.

One disease with overdominance

Susceptibility Varies with Genotype

In this section, we will examine a population of size N where there is variability in one or more parameters that affect transmission of a single pathogen. In a slight change in notation, susceptibility classes are now S_i, with baseline susceptibility σ, infectivity ϕ, and recovery/removal rate γ. Now suppose that a resistant individual's genotype reduces susceptibility by a factor σ_{ij}, so that an individual with genotype ij(where the subscipts refer to an individual with alleles i and j at a locus) has susceptibility $\sigma - \sigma_{ij}$. If we assume transmission is density dependent, such that the transmission rate is equal to βSI, then R_0 in the population is given by summing over all possible genotypes:

$$R_0 = \frac{c\phi}{\gamma}\left(S_{11}(\sigma - \sigma_{11}) + S_{12}(\sigma - \sigma_{12}) \cdots + S_{ij}(\sigma - \sigma_{ij})\right) \tag{1.6}$$

In a diploid population where there are k possible alleles, the total number of possible genotypes is equal to $\frac{k^2+k}{2}$. The size of each susceptibility class can be represented by the frequency of the genotype f_{ij} times the populations size:

$$R_0 = \frac{Nc\phi}{\gamma}\left(f_{11}(\sigma - \sigma_{11}) + f_{12}(\sigma - \sigma_{12}) \cdots + f_{kk}(\sigma - \sigma_{kk})\right)$$

Re-arranging:

$$R_0 = \frac{Nc\phi\sigma}{\gamma}\left(f_{11} + f_{12} \cdots + f_{kk}\right) - \frac{Nc\phi}{\gamma}\left(f_{11}\sigma_{11} + f_{12}\sigma_{12} \cdots + f_{kk}\sigma_{kk}\right) \tag{1.7}$$

The second half of equation (1.7) can be separated into heterozygous and homozygous genotypes:

$$\frac{Nc\phi}{\gamma}\left[\left(f_{11}\sigma_{11} + f_{22}\sigma_{22} \cdots + f_{kk}\sigma_{kk}\right) + \left(f_{12}\sigma_{12} + f_{13}\sigma_{13} \cdots + f_{kj}\sigma_{kj}\right)\right] \tag{1.8}$$

We assume neutral evolution, and that all mutations from one allelic state to another are equally likely, thus if we take the expectation with respect to genotype frequencies, all homozygous and heterozygous genotypes should have the same frequency, f_o and f_e respectively. In this case (1.8) becomes:

$$\begin{aligned} &\frac{Nc\phi}{\gamma}\left(f_o\sum_{i=1}^{k}\sigma_{ii} + f_e\sum\sum_{i\neq j}\sigma_{ij}\right) \\ &= \frac{Nc\phi}{\gamma}\left(f_o k\bar{\sigma_o} + f_e\left(\frac{k^2-k}{2}\right)\bar{\sigma_e}\right) \end{aligned} \tag{1.9}$$

Where $\bar{\sigma_o}$ and $\bar{\sigma_e}$ are the mean susceptibility for homozygotes and heterozygotes respectively:

$$\begin{aligned}\bar{\sigma_o} &= \frac{1}{k}\sum_{i=1}^{k} x_i + r x_i \\ &= \bar{x} + r\bar{x} \\ \bar{\sigma_e} &= \sum\sum_{i\neq j} x_i + x_j \\ &= 2\bar{x}\end{aligned}$$

where x_i is the contribution of allele i to the reduction in susceptibility, and r is a constant ranging from 0 to 1, representing the reduction in benefit, or redundancy, that a homozygote gets from its identical second allele.

In equation (1.9), kf_o and $(\frac{k^2-k}{2})f_e$ are equal to the total homozygosity(F) and heterozygosity respectively, as each homozygous and heterozygous genotype has the same expected frequency respectively . We define $E_f\{R_0\}$ as the expected value of R_0 taken over all possible allele frequencies in populations at mutation drift equilibrium:

$$\begin{aligned}E_f\{R_0\} &= \frac{Nc\phi\sigma}{\gamma} - \frac{Nc\phi}{\gamma}\left(F(\bar{x} + r\bar{x}) + (1-F)2\bar{x}\right) \\ &= \frac{Nc\phi\sigma}{\gamma} - \frac{Nc\phi}{\gamma}\left(2\bar{x} - F\bar{x}(1-r)\right) \\ &= \frac{Nc\phi}{\gamma}\left(\sigma - 2\bar{x} + F\bar{x}(1-r)\right) \qquad (1.10)\end{aligned}$$

If $r = 1$, then there is no advantage to increased heterozygosity. Biologically, this could be thought of as the second allele being redundant, perhaps producing addtional enzymes that provide no added benefit. If $r < 1$, then the advantage of heterozygosity goes to $\sigma - 2\bar{x}$ as F (homozygosity) goes to zero. Even when $r = 0$, ie homozygotes get no advantage from a second allele, the benefit of increased heterozygosity is more than offset by the increase in transmission resulting from an increase in population density. Furthermore, in order to benfit from heterozygosity in our model, $\bar{x}$ must be greater than zero. In otherwords, the average effect of an allele must be positive.

We can examine the change in $E_f\{R_0\}$ with respect to a change in population size N:

$$\tfrac{\mathrm{d}}{\mathrm{d}N}E_f\{R_0\} = \frac{Nc\phi}{\gamma}(\sigma - 2\bar{x} + \bar{x}(1-r)(F + N\frac{dF}{dN}))$$

Where

$$\frac{\mathrm{d}F}{\mathrm{d}N} = -\frac{4\mu(k^2+1)}{(k-1+4\mu Nk)^2}$$

We can see in Figure 1, that when $r = 1$, then the change in R_0 is constant as population size increases. When $E < 1$, R_0 increases slightly more slowly, as heterozygosity in the population increases with population size, eventually leveling out when total heterozygosity is equal to one. The rate of increase in heterozygosity will depend on the mutation rate μ and the total number of alleles k, but the decrease in R_0 due to heterozygosity will always be less than the increase due to population size.

Both Susceptibility and Infectivity Vary

If both susceptibility and infectivity vary, then equation (1.6) becomes

$$R_0 = \frac{c}{\gamma}\left(S_{11}(\sigma-\sigma_{11})(\phi-\phi_{11}) + S_{12}(\sigma-\sigma_{12})(\phi-\phi_{12})\cdots + S_{ij}(\sigma-\sigma_{ij})(\phi-\phi_{ij})\right)$$

where ϕ_{ij} it is the effect that genotype ij has on infectivity. We proceed as we did above, where $\bar{\phi_o}$ and $\bar{\phi_e}$ are the mean susceptibility for homozygotes and heterozygotes respectively. r_σ and r_ϕ represent the reduction in benefit that a homozygote gets from its identical second allele for susceptibility and infectivity, respectively.

$$\begin{aligned}
\bar{\phi_o} &= \frac{1}{k}\sum_{i=1}^{k} y_i + r_\phi y_i \\
&= \bar{y} + r\bar{y} \\
\bar{\phi_e} &= \sum\sum_{i\neq j} y_i + y_j \\
&= 2\bar{y}
\end{aligned}$$

Here, y_i is the contribution of allele i to the reduction in infectivity, and r is a constant ranging from 0 to 1. Above we assume that both susceptibility and infectivity are controlled by a single locus, and thus the index i refers to a single allele that affects both susceptibility and infectivity. We then get:

$$R_0 = \frac{Nc\sigma\phi}{\gamma} - \frac{Nc\phi}{\gamma}\left(f_o\sum_{i=1}^{k}\sigma_{ii} + f_e\sum\sum_{i\neq j}\sigma_{ij}\right) - \frac{Nc\sigma}{\gamma}\left(f_o\sum_{i=1}^{k}\phi_{ii} + f_e\sum\sum_{i\neq j}\phi_{ij}\right) + \frac{Nc}{\gamma}\left(f_o\sum_{i=1}^{k}\sigma_{ii}\phi_{ii} + f_e\sum\sum_{i\neq j}\sigma_{ij}\phi_{ij}\right)$$

Which, after taking the expectation over all populations at equilibrium, becomes

$$E_f\{R_0\} = \frac{Nc\sigma\phi}{\gamma} - \frac{Nc\phi}{\gamma}\left(2\bar{x} - F\bar{x}(1-r_\sigma)\right) - \frac{Nc\sigma}{\gamma}\left(2\bar{y} - F\bar{y}(1-r_\phi)\right) + \frac{Nc}{\gamma}\left(F(Cov(\sigma_o,\phi_o) + \bar{x}\bar{y}(1+r_\sigma+r_\phi+r_\sigma r_\phi))\right) + \frac{Nc}{\gamma}\left((1-F)(Cov(\sigma_e,\phi_e) + 4\bar{x}\bar{y})\right)$$

The two covariance terms in the above equations can be reduced to

$$\begin{aligned} Cov(\sigma_o,\phi_o) &= E(\sigma_o\phi_o) - E(\sigma_o)E(\phi_o) \\ &= (1+r_\sigma+r_\phi+r_\sigma r_\phi)\left(E(xy) - E(x)E(y)\right) \\ &= (1+r_\sigma+r_\phi+r_\sigma r_\phi)Cov(x,y) \\ Cov(\sigma_e,\phi_e) &= E(\sigma_e\phi_e) - E(\sigma_e)E(\phi_e) \\ &= 4\left(E(xy) - E(x)E(y)\right) \\ &= 4Cov(x,y) \end{aligned}$$

We define $\rho = 1+r_\sigma+r_\phi+r_\sigma r_\phi$ and end up with the follwing expression for the expected R_0 over all populations:

$$E_f\{R_0\} = \frac{Nc\sigma\phi}{\gamma} - \frac{Nc\phi}{\gamma}\left(2\bar{x} - F\bar{x}(1-\rho_\sigma)\right) - \frac{Nc\sigma}{\gamma}\left(2\bar{y} - F\bar{y}(1-\rho_\phi)\right) + \frac{Nc}{\gamma}\left(Cov(x,y) + E(x)E(y))(4 - F(4-\rho)\right)$$

As was the case when only susceptibility varied,R_0 increases with population size, but at a slight reduction in rate if $r_\sigma, r_\phi, or \rho < 1$. R_0 increases as the covariance of x and y increases, and this effect increases as heterozygosity increases.

1.4 Frequency Dependent Transmission

In all of the above calculation we have assumed density dependent transmission, where transmission is equal to βSI, such that transmission increases with an increase in population density. However, if we instead assume frequency dependent transmission, where the transmission rate is equal to $\beta\frac{SI}{N}$, and there is no increase in transmission with increasing density., so that R_0 is now equal to β. Here we assume that N is dynamic[Getz and Pickering, 1983] Under frequency dependence, the benefit of a increase in heterozygosity will not be offset by an increase in transmission rate due to increased density. If we examine the case where susceptibility varies we get an expression similar to equation (1.10), but without a factor of N:

$$E_f\{R_0\} = \frac{c\phi}{\gamma}\left(\sigma - 2\bar{x} + F\bar{x}(1-r)\right)$$

As N increases, F goes from near one to zero, and R_0 changes by a factor of $\bar{x}(1-r)$. Now, we assume that $\bar{x}$ has a maximum value of $\sigma/2$, so that no genotype can reduce susceptibility below zero. If we imagine the scenario where there are two alleles, with effects x_1 and x_2 both equal to $\sigma/2$, then the above expression reduces to:

$$E_f\{R_0\} = \frac{c\phi\sigma}{\gamma}\left(F\frac{1}{2}(1-r)\right) \tag{1.11}$$

From equation (1.11), we see that in this rather extreme scenario, $E_f\{R_0\}$ goes to zero as F goes to zero.

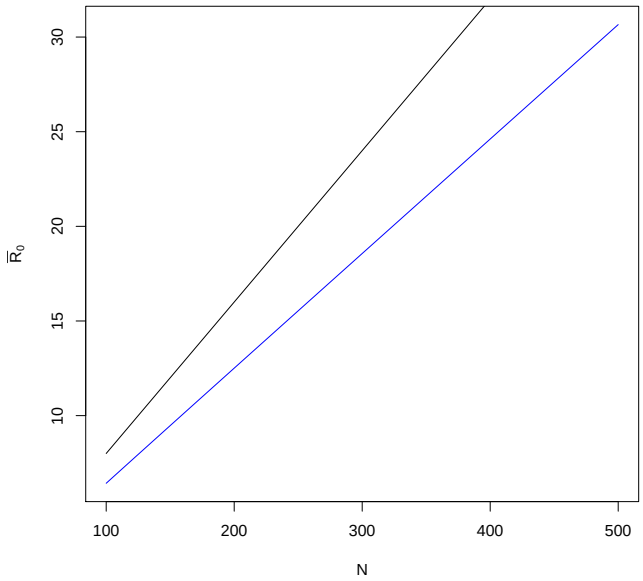


Figure 1.1: A: Expected value of $E_{D,f}\{R_0\}$ vs population size in populations exposed to multiple diseases, with a diploid population (blue, $\bar{\Delta} = 1$), and haploid (black). $\sigma = 1$, $\phi = 1$, $\gamma = 1$, $c = 0.1$, $\mu = 0.01$, $k = 50$, $z = 50$

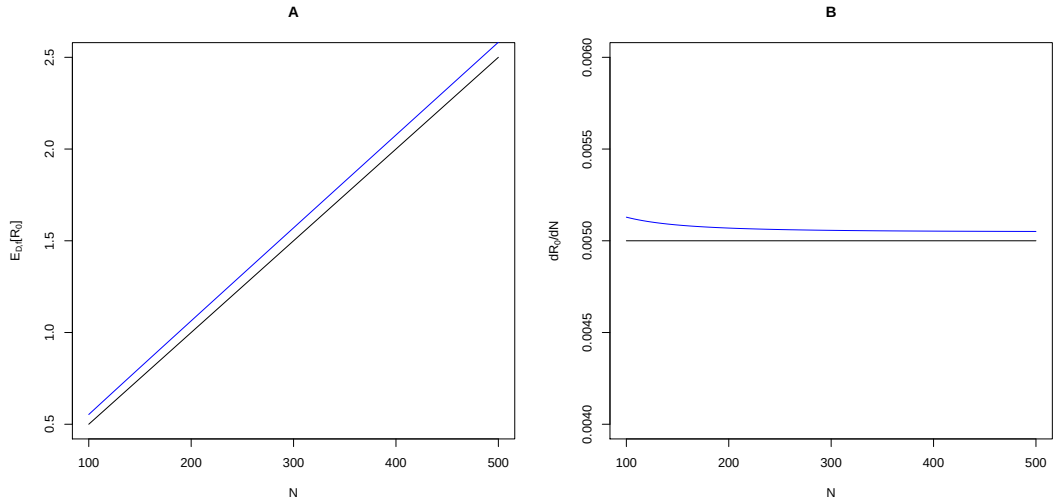


Figure 1.2: A: Expected value of$E_f\{R_0\}$ vs population size in with density dependent transmission, and heterogeneity in susceptiblilty (σ) where the second allele in a homozygote is completlely redundant(blue, $r = 0$) or effective (black, $r = 1$). B: Derivative of $E_f\{R_0\}$ with respect to change in population size. $\sigma = 1$, $\phi = 0.1$, $\gamma = 0.1$, $c = 0.01$, $\bar{x} = 0.25$, $\mu = 0.1$, $k = 50$

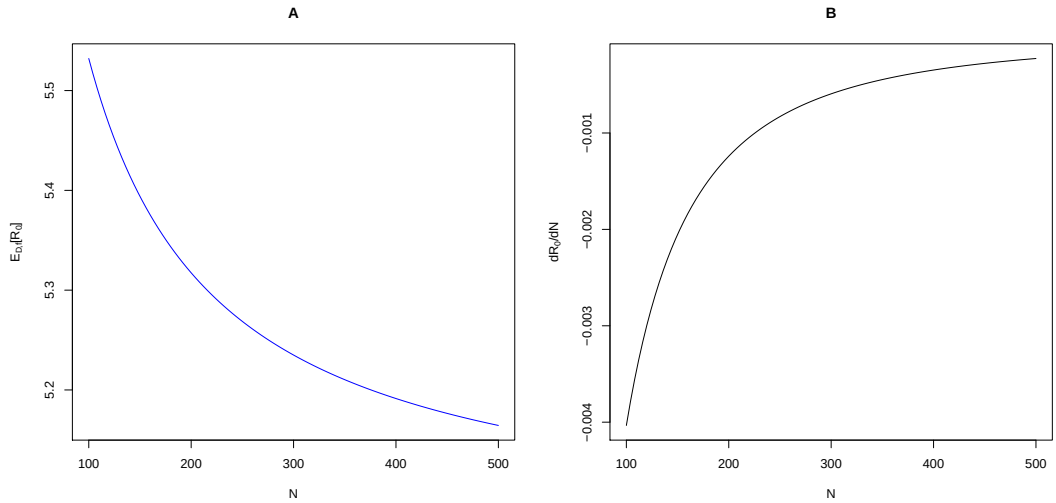


Figure 1.3: A: $E_f\{R_0\}$ vs population size with frequency dependent transmission, and heterogeneity in susceptiblilty (σ) where the second allele in a homozygote is completlely redundant(blue, $r = 0$). B: Derivative of $E_f\{R_0\}$ with respect to change in population size. $\sigma = 1$, $\phi = 0.1$, $\gamma = 0.1$, $c = 10$, $\bar{x} = 0.25$, $\mu = 0.1$, $k = 50$

1.5 Discussion

In ths paper, we hav examined how the genetic heterogeneity of a population affects the expected value of R_0. we specifically ask the question of whether increasing host population size has a beneficial effect via the increase in genetic diversity, and wheter any benefits are offset through the potentital increase in transmission rate.

Multiple Diseases

With density dependent transmission (equation (1.3)), $E_{D,f}\{R_0\}$ will always increase with increasing population size. When incorporating resistance to multiple diseases in heterozygotes, $E_{D,f}\{R_0\}$ increases slightly more slowly as population size increases (Figure 1). Under our model, heterozygous individuals benefit from two alleles, and can be resistant to two pathogens simultaneously, therefore the more heterozygous individuals in a population, the greater the overall level of resistance.

In our model we are assuming that homozygous individuals gain no additional resistance from the duplicate allele. If homozygous individuals benefit from both alleles, then there in no reduction in $E_{D,f}\{R_0\}$ due to heterogeneity. Similarly, haploid populations do not benefit from increased heterogeneity in our model (equation (1.5)).

One Disease With Overdominance

As was the case with multiple diseases we find that any reduction to $E_f\{R_0\}$ that results from an increase in genetic diversity is likely to be small (Figure 2), outside a narrow range of conditions. In the case of density dependent transmission, any reduction in $E_f\{R_0\}$ due to an increase in diversity is more than offset by the corresponding increase in population size (Figure 2). In the case of frequency dependent transmission (Figure 3), where $E_f\{R_0\}$ remains constant with increases in population size, there is a modest reduction in $E_f\{R_0\}$ due to an increase in population size and a corresponding increase in diversity. As was the case with multiple diseases, once a population is sufficiently large, any further increases in population size no longer leads to any decrease in $E_f\{R_0\}$ due to heterogeneity. Furthermore, as the benefit of the duplicate allele in a homozygote (r) increases, the reduction in $E_f\{R_0\}$ due to heterogeneity decreases, going to zero as r goes to 1.

While our study shows that the reduction of $E_f\{R_0\}$ due to heterozygosity are modest at best, it does not address any of the other potential benefits of an increase in population size. On the whole, large populations may be better off than small populations, as larger populations will be more likely to have resistant individuals that survive an outbreak of a lethal disease. Large populations will also be less susceptible to extinction due to environmental stochasticity, and allee effects at low population sizes.

In our study, we examine only the effect of heterogeneity on $E_f\{R_0\}$, and not on the probability of emergence. While probability of emergence is an increasing function of $E_f\{R_0\}$, the relationship is nonlinear [Yates et al., 2006]. While we found that an increase in heterogene-

ity has a negligible effect on $E_f\{R_0\}$, it may have a much greater effect on the probability of emergence. Yates et al. [2006] found that heterogeneity in susceptibility did not affect disease emergence, while heterogeneity in infectivity reduces the probability of emergence, which is consistent with the results of Lloyd-Smith et al. [2006].

Parameters	
σ	susceptibility
ϕ	infectivity
γ	infectious period
c	contact rate
$\bar{x}$	mean change in susceptibility due to an allele
r	effectiveness of duplicate allele
F	homozygosity
k	number of possible allelic states
μ	mutation rate
θ	$4N\mu$

Table 1.1: Parameters used in paper

Chapter 2

The effect of genetic heterogeneity on the final size of an epidemic

Introduction

When a disease invades a naive population, there is often variation in individual susceptibility, tolerance, and transmission. There may be some level of natural resistance in the population, while some individuals might spread disease more rapidly. "Superspreaders" are infected individuals who are responsible for a disproportionately large amount of the transmission [Lloyd-Smith et al., 2006], and such individuals appear to have played a significant role in the recent SARS outbreak [Shen et al., 2004]. Much of this variation is doubtless due to non-genetic factors, such as heterogeneity in contact networks, behavior, and environment. Models have shown, for example, that disease can spread more rapidly when there is heterogeneity in contact between individuals [Diekmann et al., 1990, Lloyd-Smith et al., 2006]. This paper will instead focus on how *genetic* heterogeneity may contribute to the severity of disease outbreaks. Specifically, we focus on how heterogeneity tied to genetic variation affects the final size of an epidemic, as well as how diversity does or does not persist through an outbreak.

Recently, there have been several studies examining how genetic heterogeneity contributes to an epidemic [Springbett et al., 2003, Doeschl-Wilson et al., 2011], and recent advances in sequencing techniques have shed light on how individual genes affect disease parameters [Lohmueller et al., 2003, Hirschhorn and Daly, 2005, Daetwyler et al., 2008]. There is still much to be learned about how variation at the genetic level translates to heterogeneity at the population level, and thus previous attempts to model genetic heterogeneity have taken a variety of approaches. Springbett et al. [2003] modeled heterogeneity in susceptibility, and found that genetic heterogeneity has no impact on the expected value of R_0. Doeschl-Wilson et al. [2011] compared heterogeneous populations with the same average susceptibility and found that disease risk and severity were higher in populations containing larger proportions of susceptible individuals. Lively [2010] modeled a population exposed to multiple

pathogens, each with a specific resistance genotype, and found that the average $\bar{R}_0$ is inversely related to the number of host genotypes in the population. Theoretical work by Ball [1985] demonstrated that when only susceptibility varies, increased heterogeneity decreases the final size of an epidemic.

Here, we attempt to model how host heterogeneity in susceptibility due to genetic variation affects the final size of an outbreak, as well as how that variation persists through the outbreak. We began with a theoretical approach on how variability in susceptibility affects the final size of an epidemic. Previous studies have used ad hoc distributions to generate variability in susceptibility. We attempted to use a more realistic model by conducting simulations based on Wright-Fisher evolution to investigate how genetic diversity changes before and after a disease outbreak.

Theory

The final size of an epidemic in a homogeneous population, ie a population with a single susceptibility class, can be expressed as [Diekmann and Heesterbeek, 2000]:

$$\ln \frac{S(\infty)}{S(-\infty)} = R_0 \left(\frac{S(\infty)}{S(-\infty)} - 1 \right)$$

Where $S(\infty)$ and $S(-\infty)$ are the final and starting population of susceptibles, respectively, and R_0 is the basic reproductive ratio of the population. If instead of a homogeneous population, there are different susceptibility classes, then $S(t, x)$ describes the number of susceptible individuals of class x at time t. Following Diekmann and Heesterbeek [2000], we assume that the change in the number of susceptibles during an outbreak can be described as

$$\frac{\partial S}{\partial t}(t, x) = -\frac{\mathrm{d}w}{\mathrm{d}t}(t) a(x) S(t, x)$$

Where the function $w(t)$, measures the cumulative amount of infectious materials to which individuals have been exposed at time t, and $a(x)$ is the susceptibility of an individual of type x. Then the final size of an epidemic in a heterogeneous population is:

$$S(\infty, x) = S(-\infty, x) e^{-a(x) w(\infty)}$$

$$\frac{S(\infty, x)}{S(-\infty, x)} = e^{-a(x) w(\infty)} \tag{2.1}$$

We can see from equation (2.1) that the fraction of any susceptibility class after an outbreak declines exponentially, with susceptibility (σ) times the constant $w(\infty)$ in the exponent. Again following Diekmann and Heesterbeek [2000], we can assume that the change in susceptibles can be described by

$$\frac{\partial S}{\partial t}(t,x) = S(t,x)\int_\Omega\int_0^\infty A(\tau,x,\eta)\frac{\partial S}{\partial t}(t-\tau,\eta)\,d\tau\,d\eta$$

and the final size of an epidemic can also be described by the following:

$$\ln\frac{S(\infty,x)}{S(-\infty,x)} = \int_\Omega\left(\int_0^\infty A(\tau,x,\eta)\,d\tau\right)\{S(\infty,\eta)-S(-\infty,\eta)\}\,d\eta \tag{2.2}$$

Where $A(\tau,x,\eta)$ is the force of infection acting on individuals of type x by infected individuals of type η that were infected τ units ago. Furthermore, we assume

$$A(\tau,x,\eta) = a(x)b(\eta,\tau) \tag{2.3}$$

Again, this represents separable mixing, where the force of infection acting on a susceptible individual of type x by infected individuals of type η factorizes into a product consisting of one factor representing the characteristics of the infected individual (ie susceptibility) and another factor representing the characteristics of of the infected individual (ie infectivity) and the time lapsed since infection. While there may be correlation of disease parameters within an individual, there is no correlation between individuals.

Specifically for our final size calculations, we assume that the contact rate $c = 1$ and drop it for simplicity. $a(x)$ and $b(\eta,\tau)$ are defined as

$$\begin{aligned} a(x) &= \sigma_x \\ b(\eta,\tau) &= \phi_\eta e^{-\gamma_\eta\tau} \end{aligned}$$

By combining equations (2.1), (2.2) and (2.3) (check), we get

$$w(t) = \int_0^\infty\int_\Omega b(\eta,\tau)S(-\infty,\eta)\left(1-e^{-a(\eta)w(t-\tau)}\right)\,d\eta\,d\tau$$

We take the limit $t\to\infty$ and get

$$w(\infty) = \int_\Omega\int_0^\infty b(\eta,\tau)\,d\tau\,S(-\infty,\eta)\left(1-e^{-a(\eta)w(\infty)}\right)\,d\eta \tag{2.4}$$

Which is a nonlinear scalar equation for the unknown $w(\infty)$. It can be shown that for $R_0 > 1$, equation (2.4) has a non-trivial solution. In our case, $S(-\infty,\eta)$ is just $S_i(-\infty)$, the number of susceptibles of class i. Plugging this in, along with the expressions for $a(x)$ and $b(\tau,\eta)$, we get

$$w(\infty) = \int_\Omega\int_0^\infty \phi e^{-\gamma\tau}\,d\tau\,S_i(-\infty)\left(1-e^{-\sigma_\eta w(\infty)}\right)\,d\eta$$

Integrating with respect to τ we get

$$w(\infty) = \frac{\phi}{\gamma} \int_\Omega S_i(-\infty) \left(1 - e^{-\sigma_\eta w(\infty)}\right) \, d\eta$$

We assume susceptibles form discrete classes, each with susceptible σ_i for $i \dots k$. The above integral becomes the sum over all possible susceptibility classes:

$$w(\infty) = \frac{\phi}{\gamma} \sum_{i=1}^{k} S_i(-\infty) \left(1 - e^{-\sigma_i w(\infty)}\right) \tag{2.5}$$

$$= \frac{\phi}{\gamma} \left(N - \sum_{i=1}^{k} S_i(-\infty) e^{-\sigma_i w(\infty)} \right) \tag{2.6}$$

$$= \frac{N\phi}{\gamma} - \frac{\phi}{\gamma} \sum_{i=1}^{k} S_i(-\infty) e^{-\sigma_i w(\infty)} \tag{2.7}$$

more generally, when several parameters vary by genotype:

$$\begin{aligned} w(\infty) &= \sum_{i=1}^{k} \frac{\phi_i}{\gamma_i} S_i(-\infty) \left(1 - e^{-\sigma_i w(\infty)}\right) \, d\eta \\ &= \sum_{i=1}^{k} \frac{\phi_i}{\gamma_i} S_i(-\infty) - \sum_{i=1}^{k} \frac{\phi_i}{\gamma_i} S_i(-\infty) e^{-\sigma_i w(\infty)} \\ &= N \cdot E\left(\frac{\phi}{\gamma}\right) - \sum_{i=1}^{k} \frac{\phi_i}{\gamma_i} S_i(-\infty) e^{-\sigma_i w(\infty)} \end{aligned}$$

If only susceptibility varies, then the RHS of equation (2.7) reduces to moment generating function of σ, $M_\sigma(t)$, with $t = -w(\infty)$:

$$w(\infty) = \frac{N\phi}{\gamma} - \frac{\phi}{\gamma} \sum_{i=1}^{k} S_i(-\infty) e^{-\sigma_i w(\infty)} \tag{2.8}$$

$$= \frac{N\phi}{\gamma} \left(1 - \sum_{i=1}^{k} \frac{S_i(-\infty)}{N} e^{-\sigma_i w(\infty)} \right) \tag{2.9}$$

$$= \frac{N\phi}{\gamma} \left(1 - E(e^{-\sigma w(\infty)})\right) \tag{2.10}$$

$$= \frac{N\phi}{\gamma} \left(1 - M_\sigma(-w(\infty))\right) \tag{2.11}$$

$$M_\sigma(-w(\infty)) = 1 - w(\infty)m_1 + \frac{w(\infty)^2 m_2}{2!} - \frac{w(\infty)^3 m_3}{3!} + \cdots \tag{2.12}$$

Where $m_1, m_2, ...m_i$ are the ith moment of the distribution determining the frequency susceptibles. Increasing m_2 ($E(\sigma^2)$) will decrease the solution to equation (2.9) when other moments are held constant.

Simulations

Genetic Model

We conducted simulations to determine the effect of genetic diversity on both the final size of an epidemic and the amount post epidemic standing diversity. To simulate pre-epidemic diversity, we simulated evolution for 10,000 generations. We assume mutations between alleles are specified by a k-allele model, where the mutation rate between alleles is specified by a parameter μ, and an allele i mutates into each of the $k-1$ other allelic states with equal probability. This results in a k-allele Wright-Fisher model [Ewens, 2004]. We assume that the parameters are determined by a single locus, and any covariance in parameters is due to effects at the single locus. We assumed that when a population is free of disease, any variation in the population is maintained through neutral evolution. Though our model assumes a haploid population, it is easily extended to a diploid population.

In order to simulate a range of genetic diversity for a given population size, we also used a coalescent model to simulate ten fold reduction in population size, followed by an immediate return to the starting population size. Specifically, if our starting population is of size N, we used a coalscent model to draw a sample of size $N/10$ and used these gene frequencies as the starting population for two generations of Wright Fisher evolution of a population of size N.

Disease Model

Using gene frequencies obtained as above, we then simulated an epidemic using a modification of the basic SIR (Susceptible, Infected, Recovered/Removed) compartmental model frequently used to describe infectious disease dynamics [Anderson and May, 1991]:

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= -\beta SI \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \beta SI - \gamma I \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma I\end{aligned}$$

The above model describes an epidemic of a directly transmitted disease through a population in which S, I and R are the numbers of susceptible, infected and resistant individuals, respectively. The model assumes that the total population size $N = S + I + R$, where β is the transmission rate between infected and susceptibles, and γ is rate of recovery into the resistant/removed class. The model describes only short term dynamics, and does not

include births or deaths. The basic reproductive ratio R_0, using the above model, is $\beta N/\gamma$. In our model we assume that the susceptible class can be divided into k subgroups, each labeled S_i. We break down the transmission process β into infectivity ϕ, susceptibility σ, and contact rate c, and consider that each of these parameters could potentially vary by genotype:

$$\begin{aligned}\frac{\mathrm{d}S_i}{\mathrm{d}t} &= -\sum_j \sigma_i \phi_j c_{ij} S_i I_j \\ \frac{\mathrm{d}I_i}{\mathrm{d}t} &= \sum_j \sigma_i \phi_j c_{ij} S_i I_j - \gamma_i I_i \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \sum_i \gamma_i I_i\end{aligned}$$

We define susceptibility as the probability of becoming infected given an infectious contact with an infected individual, and infectivity as the probability that a contact is potentially infectious, given a contact with a susceptible. Full contact between groups was assumed with random mixing of all individuals, and we assume separable mixing, ie the absence of correlation between infectivity and susceptibility between individuals (as opposed to within individuals) [Diekmann et al., 1990]. In other words, transmission between a susceptible individual of type i with properties a_i and an infected individual of type j with properties b_j is equal to $a_i b_j$. In a population with separable mixing, R_0 is just the average R_0 over all sub-populations $\sum \beta_i N_i/\gamma_i$. In the model below, we will consider both density dependent and frequency dependent transmission. For the purposes of this paper, we will define resistance to a disease as a change in any of the parameters that affect transmission, such as a decrease in susceptibility σ or an increase in the recovery rate γ, for example.

In our model, susceptibility σ takes on k values, each corresponding to an allele in the genetic model, and ranges in value from near 0 to 1. Each realization of our simulations result in a sample of n out of k alleles, and therefore a system of $3n$ differential equations describing the disease dynamics. These equations were solved using the deSolve package in the program R [R Core Team, 2012].

We ran a total of 10,000 simulations, with 5,000 each of bottlenecked and non-bottlenecked populations. Disease parameters were set so that the expected value of R_0 was approximately 5. See Table 1 for all parameters used in the simulations and their values.

Diversity Metrics

We measured genetic diversity using three different metrics. We used allelic richness, or number different types in the population, the Simpson index, and the Shannon index. The Simpson index is given by:

$$\sum_{i=1}^{n} p_i^2$$

Where n is the number of alleles in the sample, and p_i is the frequency of the ith allele. The Simpson index ranges from $1/n$ when all allele frequencies are identical, to 1 when there is a single dominant allele. In a diploid population, the Simpson index is equal to the expected homozygosity.

The Shannon index is given by:

$$-\sum_{i=1}^{n} p_i \log p_i$$

with n and p_i as described above. The Shannon index is equal to $\log n$ when all allele frequencies are equal, and near zero when there is one dominant allele. Both the Simpson and Shannon indices indicate the evenness of the representation of alleles in the population.

For reach realization of the simulations, we recorded the values of the three diversity indices in the initial population for both the pre- and post-epidemic population. For each index, we calculated the change in diversity as:

$$\frac{initial\,value - final\,value}{final\,value}$$

We also recorded the final size of the epidemic,R_0, and the mean and variance of the susceptibility for both the pre- and post-epidemic populations. Results were fit using the package in R [R Core Team, 2012] and graphs of the simulation results were constructed using the R package [Wickham, 2009]

Simulation Results

In our simulations diversity tended not to persist through an outbreak (Fig 3). Populations with low allelic diversity pre-epidemic had similar post-epidemic allelic diversity and evenness compared to populations with high pre-epidemic allelic diversity. Populations with high and low pre-epidemic evenness had similar changes in post-epidemic allelic diversity. Evenness was reduced in post-epidemic populations. This is most apparent in the change in the Simpson index, while the response in the other indices was essential flat. Not surprisingly, there was more variability in the change in pre to post-epidemic diversity when the pre-epidemic population had higher diversity and evenness. This is likely because a populations with low diversity are bounded in the potential decrease in diversity.

Figure 4 depicts the final size of the epidemic as a function of both the Simpson and Shannon index. Both plots exhibit some curvature, with lower evenness associated with both large and small epidemics. As expected, low evenness and a high mean susceptibility are associated with large epidemics, and low evenness and a low mean susceptibility are associated with smaller epidemics.

Fig 5 shows the final size of an epidemic plotted against the variance of susceptibility in the pre-epidemic population. As predicted from (2.9), the final size of the epidemic decreases as the variance in susceptibility in the pre-epidemic population increases. Though the effect is rather modest, when the is split based on the median susceptibility, the effect increases (Fig 6).

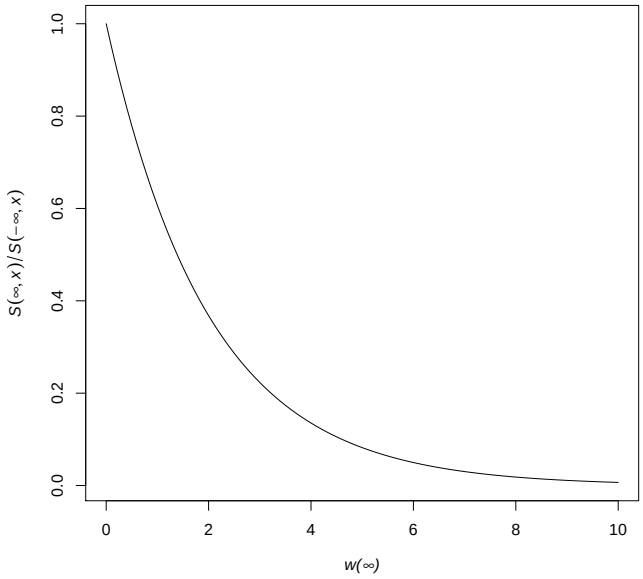


Figure 2.1: Final size of an epidemic in susceptible population x (equation (2.1)) plotted against $w(\infty)$

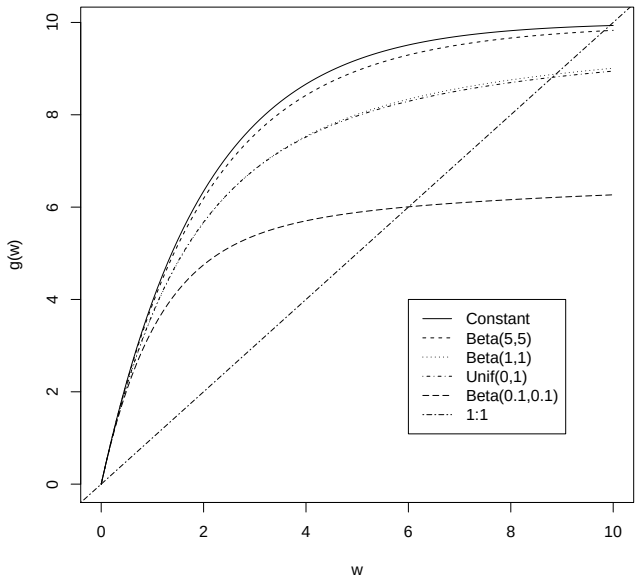


Figure 2.2: Equation (2.9) (here noted as $g(w)$) plotted as a function of w, with σ having the same mean but different levels of variance. The solution to the scalar equation in (2.9) is represented by the intersection with the $1:1$ line. As variance increases, the solution to (2.9) becomes smaller. $\phi = 0.01$, $\gamma = 1$, $c = 1$, $N = 1000$

Parameters		Value
σ	susceptibility	$(0.001, 1)$ in $K-1$ equally spaced intervals
ϕ	infectivity	0.01
γ	infectious period	1
μ	mutation rate	0.001
K	number of alleles	50
N	population size	1000
B	size of population bottleneck	0.1

Table 2.1: Parameters and their values in the simulations

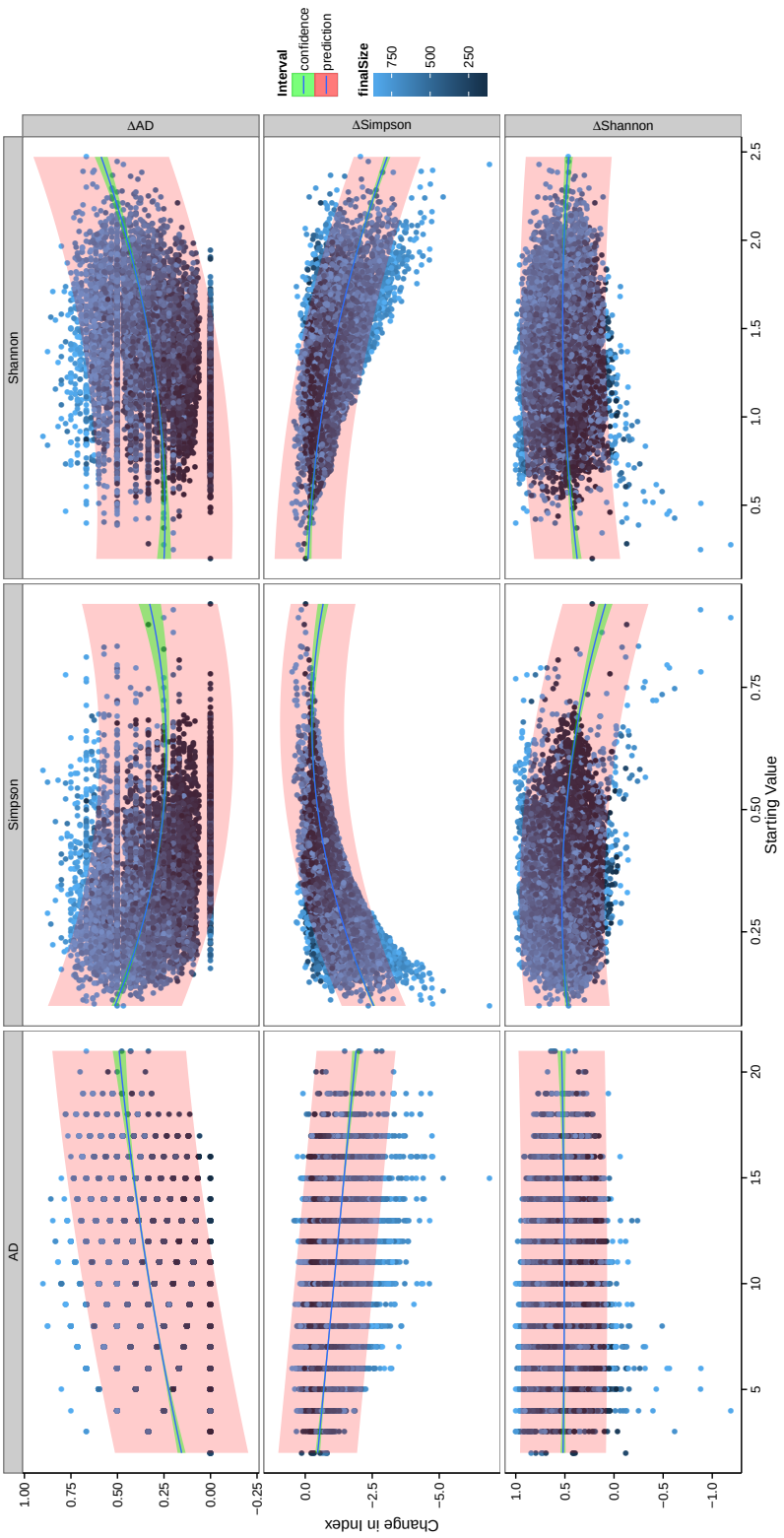


Figure 2.3: Change in each index plotted against initial value for 10000 simulations. The color of each point represents the final size of the epidemic, where lighter colors represented larger epidemics, and lighter colors representing smaller epidemics. Fits was done with a quadratic function, with point wise 95% confidence interval in red and prediction interval in green.

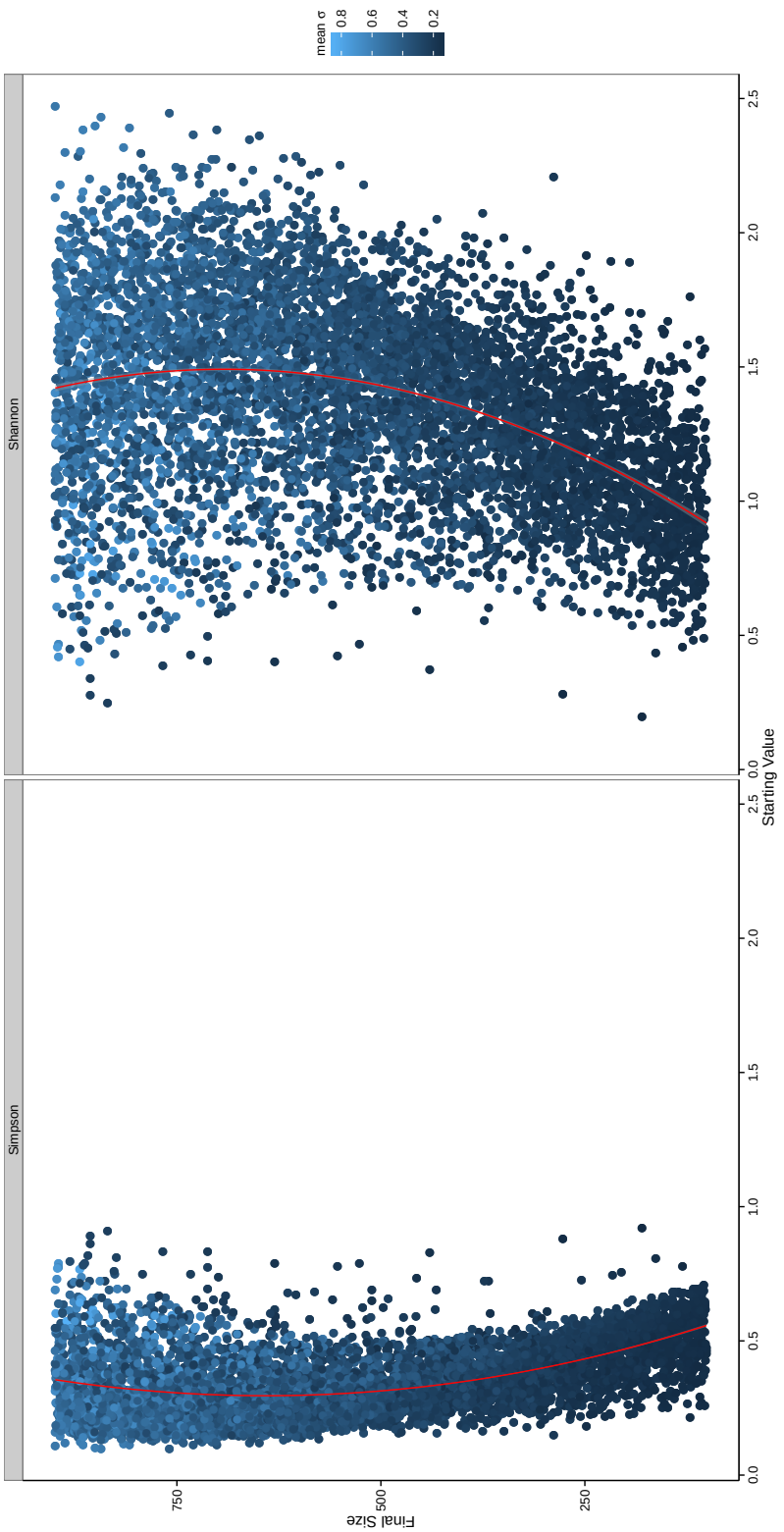


Figure 2.4: Final size of the epidemic, plotted against both Simpson and Shannon indices for the pre-epidemic population, where color indicates mean susceptibility σ of a pre epidemic population. Red lines indicate a quadratic fit to the data with the axes flipped

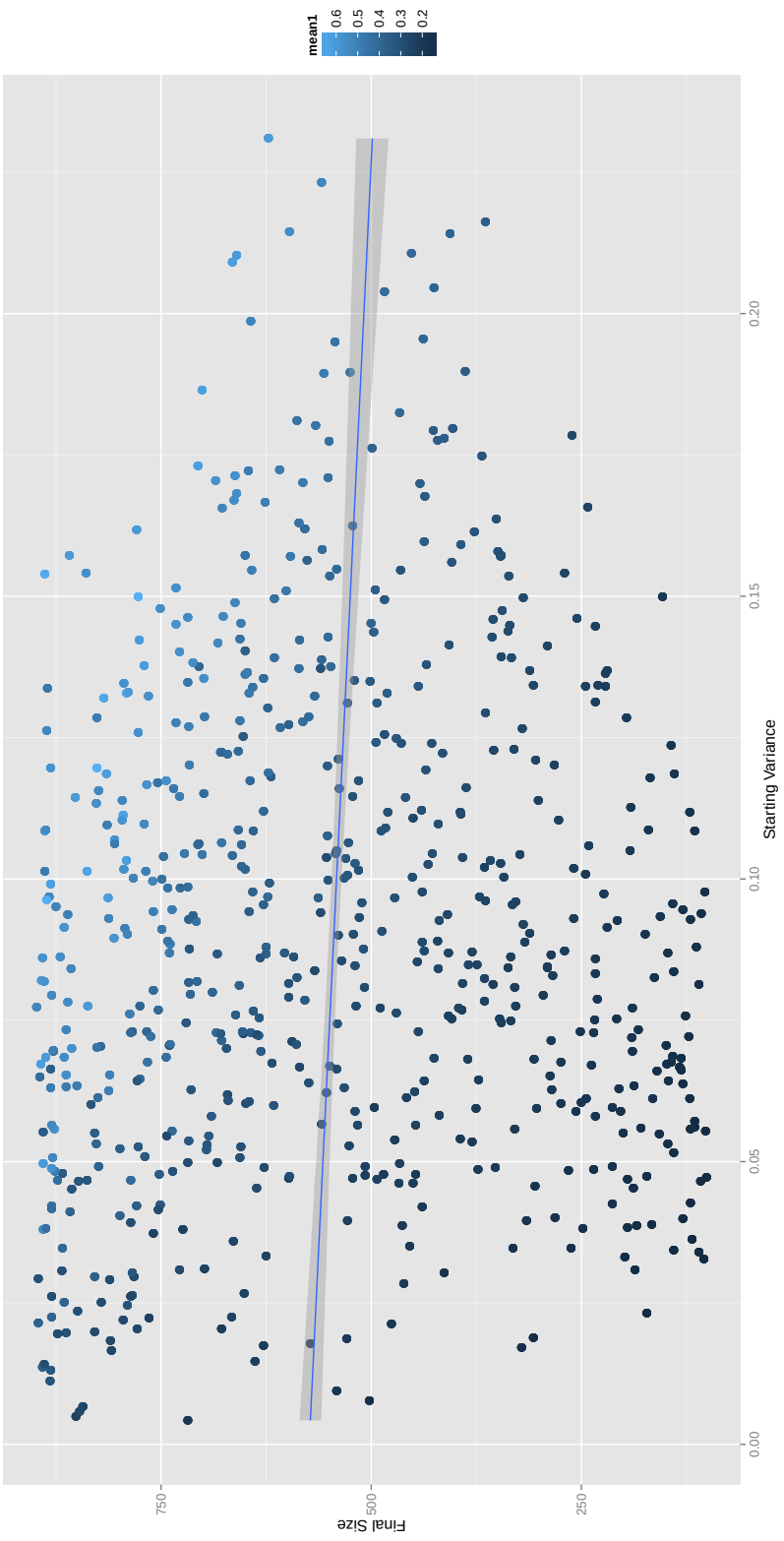


Figure 2.5: Final size of an epidemic in plotted against variance in susceptibility σ. Color corresponds to mean susceptibility. Estimated slope of -325 $\pm$ 67

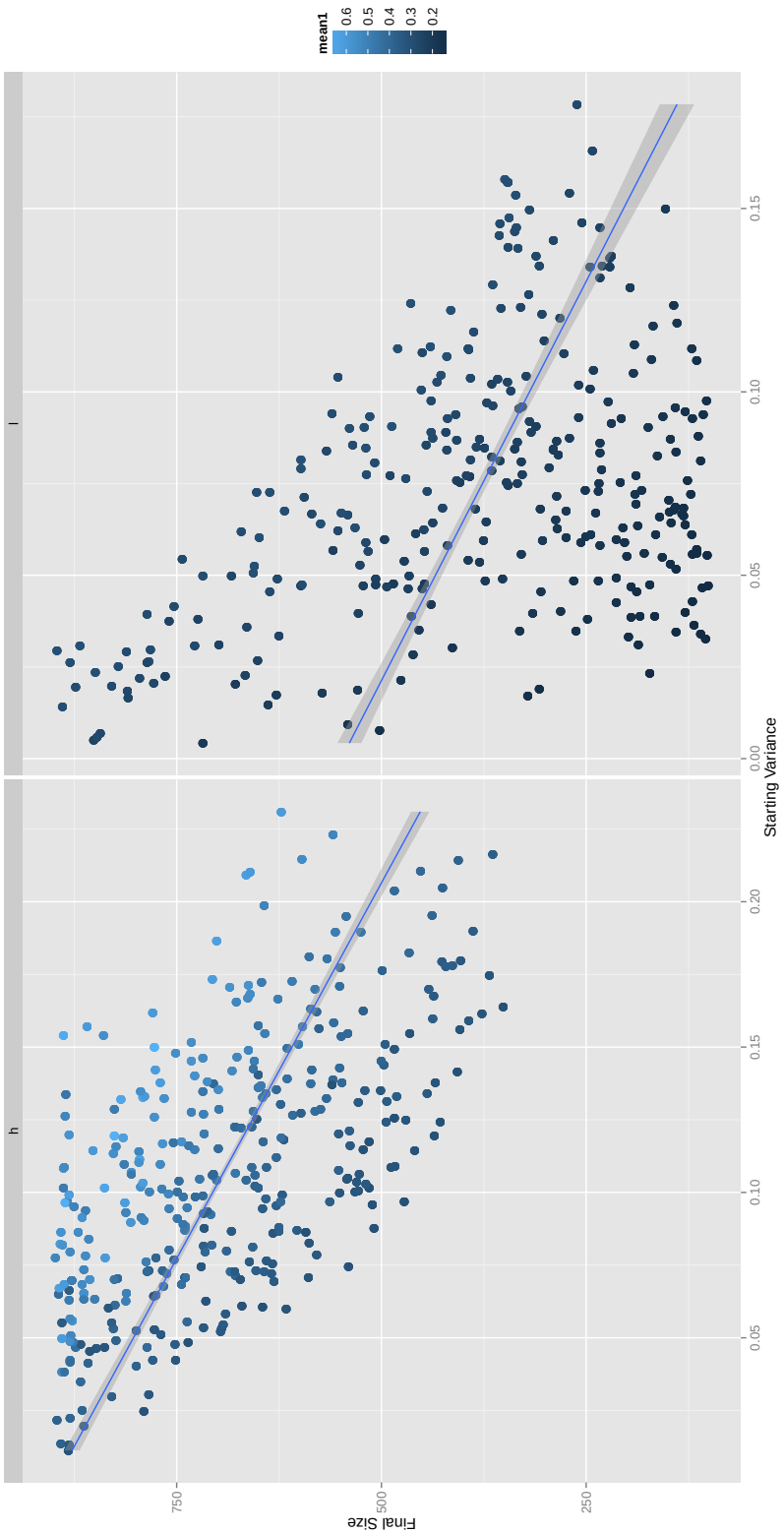


Figure 2.6: Final size of an epidemic in plotted against variance in susceptibility σ with data split in two based on median of the mean susceptibility. Color corresponds to mean susceptibility. Estimated slope of -1935 $\pm$ 44 for left panel, -2297 $\pm$ 96 for right panel.

Discussion

Effect of Diversity on the Final Size of an Epidemic

We can see from equation (2.1) that the fraction of any susceptibility class after an outbreak declines exponentially (Figure 4), where the value of $w(\infty)$ is given by the nonlinear scalar equation in (2.9). We can see from equation (2.9) that $w(\infty)$ has an upper bound of $N \cdot E(\phi/\gamma)$. The larger the value of $w(\infty)$, the steeper the decline in susceptibles. As shown in equation (2.12), when only susceptibility varies, the value of $w(\infty)$ is determined largely by the moment generating function for σ , with $w(\infty)$ increasing as $E(\sigma^2)$ increases (Figure 5). This is consistent with Ball [1985], who found that when only susceptibility varies, increased heterogeneity leads to smaller outbreaks. Our results imply that higher moments will also have an effect on the final size, given by (2.12).

Since, in our model, the number of survivors declines exponentially, small changes in susceptibility lead to large changes in survivorship after an outbreak. We also find that the frequency of any genotypes other than those corresponding to the most resistant individuals are likely to be severely reduced following an outbreak with high mortality. While genetic diversity may persist though the outbreak, the frequency spectrum will be shifted toward resistant genotypes. This was verified by our simulation study, as both evenness and diversity declined post outbreak. This is to be expected, as our scenario represented a strong selective event. Our simulations also suggest that the final size of an epidemic is a bifurcating function of the evenness of a population, with populations with low mean susceptibility and low evenness undergoing small or no epidemics, and populations with high mean susceptibility and low evenness experiencing larger epidemics. This is to be expected, as resistance is maximized when there is a single resistant allele in the population, and minimized when there is a single, highly susceptible genotype. However, for populations with similar mean susceptibilities, more diverse populations will have smaller epidemics.

There are several ways our approach could be extended. In our simulations, we assumed all populations were haploid. A diploid model could incorporate phenomena such as overdominance, where heterozygotes have lower susceptibility on average than homozygotes. This could allow genetic diversity to have a greater effect on the final size of an epidemic. A diploid model could also incorporate change in expected homozygosity after an epidemic. Our model for susceptibility is almost certainly overly simplistic. Resistance to a disease is likely the result of the interactions of many different genes. We focused on simplicity in our study, but a more complex mutation model incorporating multiple genes could yield more realistic results.

In our simulations we only tested the effect of variance in susceptibility on the final size of an epidemic. (2.12) indicates that higher moments will have an effect on the final size as well. While we only addressed heterogeneity in susceptibility in this study, it would be possible to extend our modeling and simulation framework to variation in other parameters such as infectivity, recovery time, and contact rate. However, because of the form of equation (2.9), we could not use the moment generating function approach used in (2.12). A different

theoretical or simulation based approach would have to be used to address variation in other parameters.

Both our theoretical model and simulations show that while more heterogeneous populations have smaller epidemics on average, the evenness of the diversity of a populations declines markedly. While there may be small change in allelic diversity post-epidemic, as some alleles remain in the population at low frequencies, epidemics result in large shifts in the frequency spectrum towards less susceptible alleles.

Chapter 3

Genetic diversity in the *Rana muscosa/Rana sierrae* species complex

Introduction

California's mountain yellow-legged frog (a species complex consisting of *Rana muscosa* in the southern part of its range and *Rana sierrae* in the north), has declined dramatically in the last 50 years, with declines of over 90% across its range [Vredenburg et al., 2007]. The decline of *R. muscosa* in southern California has been even more severe, with extinction at >99% of historical sites [Backlin et al., 2004]. These highly aquatic species historically occurred in both lake and stream habitat in the high elevations (>2000 meters) of the Sierra Nevada Mountains in California [Grinnell and Storer, 1924]. The introduction of non-native trout has vastly reduced and fragmented populations throughout their range [Vredenburg et al., 2007], and the introduction of the fungal pathogen *Batrachochytrium dendrobatidis* (hereafter referred to as Bd) has all but driven the remaining populations to extinction [Briggs et al., 2005].

Bd has caused declines in amphibian populations throughout the globe. Bd is highly pathogenic in many species, with low levels of initial infection leading to high levels of mortality [Skerratt et al., 2007], while other species appear to tolerate Bd infection with little to no side effects [Weldon et al., 2004, Schloegel et al., 2010] and may serve as reservoirs for other species. While most *R. muscosa/sierrae* populations exposed to Bd appear to either go extinct or undergo dramatic declines [Vredenburg et al., 2010], there do appear to be a small number of *R. muscosa* and *R. sierrae* populations that are persisting with Bd [Briggs et al., 2010]. These 'persistent' populations have tested positive for Bd for as long as they have been surveyed, is some cases going back 10 years or more. These populations are often much smaller and less dense than disease-free populations. The earliest record of Bd infection in California may be in a *R. muscosa* misidentified as a *R. boylii*, indicating that Bd may have

been present in some populations since the 1960s [Vredenburg et al., 2010].

Several studies have suggested that reduced genetic diversity or 'genetic erosion' can lead to reduced fitness and make resistance to pathogens more demographically costly [Luquet et al., 2012, 2011]. In this study, we use microsatellite genetic data to determine if *R. muscosa/sierrae* have reduced levels of genetic diversity. Furthermore, we examine populations that have been persisting with Bd, to determine if they have signatures of genetic bottlenecks that may have been the result of recent disease induced mortality similar to that observed in outbreak populations.

Materials and Methods

Field Sampling

Rana muscosa Populations

Unites States Geological Survey personnel collected tissue samples of *Rana muscosa* during surveys in southern California in 2003-2009 from nine unique locations: South Fork Big Rock Creek, Little Rock Creek, Bear Gulch, Vincent Gulch, Devils Canyon, East Fork City Creek, Fuller-Mill Creek, Dark Canyon, Tahquitz Creek, and one captive population of individuals originally collected in Dark Canyon. Additionally, we obtained samples from the nearest extant population in the southern Sierra Nevada, Milestone Basin, in Sequoia Kings National Park, in 2004. Samples in Milestone Basin were collected during a large Bd die-off described in [Vredenburg et al., 2010]. Other populations in the southern Sierra Nevada are considered extinct [Vredenburg et al., 2007]. Tissue consisted of toe clips of adult and sub-adult (post-metamorphic) frogs, and tail clips of tadpoles, which we preserved in 95% ethanol until extraction. Sample size ranged from 14 in the Vincent Gulch population to 91 in the East Fork City Creek population, for a total of 614 individuals across all populations (Table 1), representing the majority of animals alive and present in southern California for the last decade. All sampled locations are shown in Figure 1.

We also included data from Bingham [2007], which were collected from three large, pre-decline populations in 60 Lake Basin, located in Sequoia Kings National Park, using the same loci with consistent scoring.

Rana sierrae Populations

We obtained samples from Conness Pond in Yosemite National Park in 2003 and Ebbetts Pass in El Dorado National Forrest in 2004. Together with Unites States Geological Survey personnel, we also collected tissue samples of *Rana sierrae* in 2003-2005 in El Dorado National Forrest from four locations: Aloha Pond, Hell Hole, Waca Pond, and Pyramid Lake. Sample size ranged from 5 individuals in the Aloha population to 48 in the Conness Pond population.

Real Time PCR

Before tissue was collected, each animal swabbed to quantify the amount of Bd present on the individual by rubbing a sterile, cotton swab over the ventral surfaces and digits [Hyatt et al., 2007, Vredenburg et al., 2010]. Swabs were extracted and analyzed on a ABI StepOne Plus real time pcr machine following the protocols of Boyle et al. [2004]. All Southern California populations were swabbed, as well as those at Ebbetts Pass and the Tahoe Basin populations. The Conness Pond population was Bd positive, as it has persisted with Bd since at least 2001 (R. Knapp unpublished).

Genetic data collection

Genomic DNA was extracted using PrepMan Ultra^{TM} reagent (Applied Biosystems), with the following modifications to the manufacturer's protocol: tissue was combined with 45 ll of PrepMan and 40 mg of silica beads, shaken on a beadbeater for 90s, digested and centrifuged for 30 s. We amplified each individual at nine microsatellite loci by polymerase chain reaction (PCR) using primers designed by Genetic Information Services (GIS). PCR products were run on a 3730 capillary sequencer with the GeneScan Liz 500 size standard (Applied Biosystems). Primer sequences are shown in Figure 4.

Population Genetic Analysis

Population structure within a sampling unit can lead to false signatures of population bottlenecks [Chikhi et al., November 2010], therefore population structure in the microsatellite data was assessed using a Bayesian clustering algorithm implemented in the program STRUCTURE v2.3.4, with population source used as prior information (Hubisz et al., 2009; Pritchard et al., 2000). We used the admixture model with correlated allele frequencies to account for any migrants in the dataset, following recommendations of Francois and Durand (2010). STRUCTURE was run for the adult frog dataset by setting the cluster ('k') value incrementally from 1 to the maximum number of sampled units (lakes) with 10 independent runs at each k value. A burn-in period of 100,000 steps was followed by MCMC sampling for 1 million steps. Results from this analysis were used to determine sampling units for the bottleneck analysis.

Variability at each microsatellite locus was tested for deviation from Hardy-Weinberg equilibrium (HWE) using chi-square and Fisher's Exact tests. This was conducted within each population for all available samples and for a subset of the data including only adult frogs. Due to the large number of tests, the level of statistical significance ($\alpha = 0.05$) was adjusted by Dunn-Sidak correction $(1-(1-\alpha))^{1/n}$. The mean number of alleles (allelic richness), observed and expected heterozygosity, and fixation index (F_{is}) were averaged across loci for each population. Statistical significance was assessed by permuting the data 10000 times. All statistical tests were conducted within R and ARLEQUIN v3.5.1.2 (Excoffier and Lischer 2010).

To test for genetic bottlenecks in each population, we calculated the sign-test [Cornuet and Luikart, 1996] and M-ratio statistic [Garza and Williamson, 2001]. The sign-test examines whether there is an excess in expected heterozygosity across loci, which occurs when the effective population size is sharply reduced during a population bottleneck. This was calculated in the program BOTTLENECK (Piry et al. 1999) assuming the two-phase model (T.P.M.) with variance set to 30 and probability set to 70%, and significance assessed over 10 000 replicates. The M-ratio statistic examines the ratio of the number of alleles and the allele range, with the expectation that the number of alleles declines faster than the allele range in a bottlenecked population. The M-ratio was calculated in ARLEQUIN and tested for significance by comparison to simulated data. Using R, we simulated genetic diversity in a population with constant size at a microsatellite locus evolving under two-phase mutation model, with a mean size of multistep mutation set to 3.5 and the proportion of single steps set at 0.89 (as suggested by Garza and Williamson 2001, based on their survey of the literature). From each simulated dataset, a sample was drawn corresponding to the number individuals in the population sample, from a total of 10 000 datasets. Critical values ($\alpha = 0.05$) of the M-ratio, below which an M-ratio value would be likely to indicate a bottleneck, were determined for two levels of ancestral theta, $\theta = 1$ and 10. For comparison across populations, we report critical values for a sample of 20 individuals. This represents the lower end of our sampling numbers, and is therefore a conservative threshold for most of the populations used in the analysis.

Results

Real Time PCR

Of the *Rana musocsa* populations tested, only the Tahoe Basin came up positive for Bd, and was positive over all 3 years sampled (Table 1), with a prevalence of 0.87 at Hell Hole, 0.62 at Pyramid Lake, and 0.29 at Waca Pond.

Genetic Variation

Several populations exhibit statistically significant deviations from HWE after Dunn-Sidak correction, although the number of significant tests is higher using the chi-square method versus Fisher's Exact method. When only adult frogs are considered, the number of significant tests declines, irrespective of the statistical method. There is no clear trend of deviation from HWE at specific loci across all populations, which would otherwise suggest the presence of null alleles. Similarly, there is no clear predominance of HWE deviations in specific populations across all microsatellite loci, which suggests age stratification or admixture are not major factors. As a result, all downstream analyses included all available data.

The analysis of microsatellite variation within *R. sierrae* using STRUCTURE identified Conness Pond, Ebbett's Pass, and the Tahoe Basin as distinct populations, with a plateau

in the log likelihood of the data at k=3. There was not significant support for structure within the Tahoe Basin (Fig 2), and therefore this was treated as one population. For the Milestone and Southern California *Rana muscosa* populations STRUCTURE shows a distinct plateau in the log likelihood of the data at nine clusters. The eight clusters that are strongly supported represent every distinct sampling site, with the exception that Vincent Gulch and Bear Gulch are represented as one cluster, and the sample from Tahquitz Creek and the captive population are not distinguished from the resident Dark Canyon population. There support for structure within Milestone Basin was weak, and it was treated as a single population for the rest of the analysis.

Comparison of microsatellite genetic variability averaged across loci (Table 1) show the number of alleles and observed heterozygosity are similar across populations. Differences in the observed and expected heterozygosity, as measured in the fixation index, show that most populations do not have a significant deficit of heterozygous individuals. However, East Fork City Creek, Little Rock Creek, and Ebbetts Pass do have deficits that differ significantly from zero, although outside of the Ebbetts Pass population, the level of inbreeding is not high. Several populations, including Vincent Gulch, Bear Gulch, Fuller-Mill Creek and the captive population have a larger observed versus expected heterozygosity, resulting in negative fixation indices, but these are not significantly different from zero.

Simulated critical M-values varied by sample size, so we used used the lowest vales for each choice of θ ($\theta = 1$ and $\theta = 10$). This corresponds to a somewhat conservative critical vale due to the small populations samples in some of the populations. Using a M-crit of 0.41 and 0.42 for ($\theta = 1$) and ($\theta = 10$) respectively, there are signatures of population bottlenecks across southern California *R. muscosa* populations (Table 2). Based on the sign-test, only the Dark Canyon population shows a significant ($p = .0005$) bottleneck effect. Using the M-ratio statistic, however, all Southern California populations have values below critical thresholds of a sample drawn from a moderate ($\theta = 1$) to large ($\theta = 10$) population. Only the Milestone Basin population, sampled during a disease outbreak, showed no sign of a bottleneck, although the M-value of 0.534 was close to the critical simulated value.

None of the populations from 60 Lake Basin or any of the *R. sierrae* populations had M-values indicating population bottlenecks.

Discussion

Genetic diversity, as measured by mean number of alleles per locus and expected heterozygosity, is quite low compared to related frogs in the genus *Rana* [Monsen and Blouin, 2004, Zhan et al., 2009, Zhao et al., 2009]. Similarly low levels of microsatellite variation are evident in other threatened or endangered ranids, including the Columbia spotted frog, *Rana luteiventris* [Funk et al., 2005], the Italian agile frog, *R. latastei* [Ficetola et al., 2007], and the northern leopard frog, *R. pipiens* [Wilson et al., 2008]. Despite the reduced levels of genetic variability, levels of inbreeding, as measured by F_{is} are currently low in most populations of *R. muscosa/sierrae*, with the highest values found in the East Fork City Creek,

Little Rock Creek, South Fork Big Rock Creek and Milestone Basin populations (F ranges from 0.1 to 0.26). Although populations of southern California *R. muscosa* are not inbred, there is evidence that genetic diversity has recently declined.

Population bottleneck tests demonstrate that several populations have had a significant reduction in allelic diversity relative to the range of alleles. The M-ratio tests show significant population size reduction in all of the southern populations, while the sign-test only shows a significant result in the Dark Canyon population. It is well known that there is less statistical power to detect a bottleneck with the sign-test [Williamson-Natesan, 2005], which is based on measuring an excess of observed heterozygosity relative to the expected heterozygosity (calculated from the observed number of alleles). The negative fixation indices in several additional populations (Table 1) show a trend towards an excess of observed heterozygosity, but unless a strong bottleneck is very recent and/or ongoing, heterozygosity is expected to rapidly return to equilibrium values.

Although we were not able to sample enough persistent populations to determine if there were differences in diversity between populations persisting with Bd and those not yet exposed, neither of the two persistent populations sampled (Conness and Tahoe Basin) showed any indication of population bottlenecks. This suggests that, at least in these populations, the initial invasion of Bd was not accompanied by large die-offs. More persistent populations should be sampled to see if this pattern holds, however, our work indicates that invasion of Bd does not always result in a severe population crash. The populations that did show bottleneck signature were all in Southern California, which is consistent with this area experiencing the steepest declines in population, largely due to habitat loss and fragmentation [Backlin et al., 2004]. These populations all tested negative for Bd at the time of our study; it remains to be seen if the lack of genetic diversity will result in more severe outbreaks when Bd finally reaches these populations.

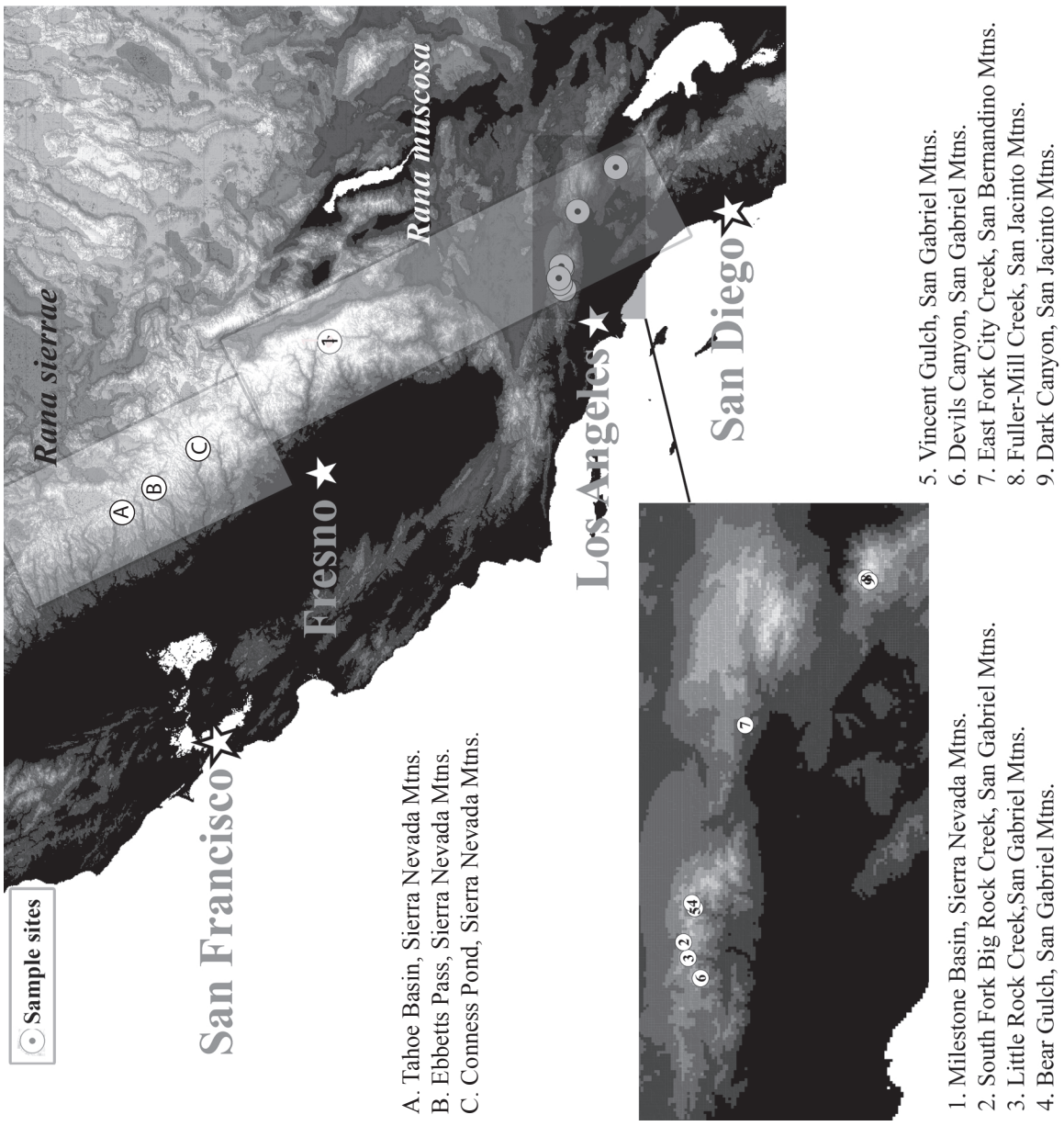


Figure 3.1: Map of localities used in this study. Sites A-C correspond to *R. sierrae* populations. Sites 1-9 correspond to *R. muscosa*

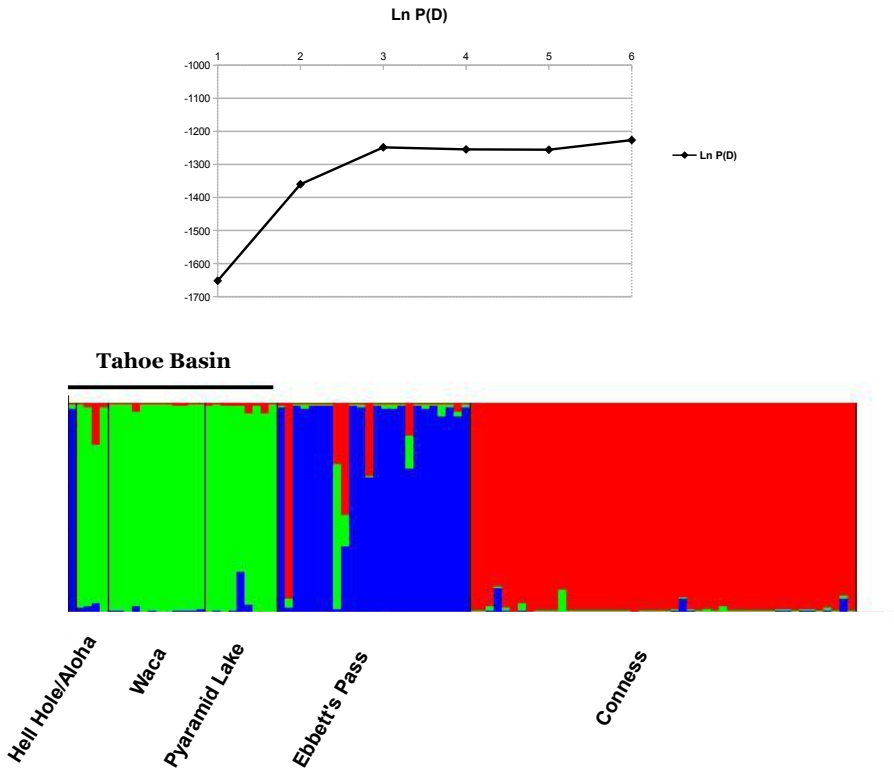


Figure 3.2: STRUCTURE analysis of microsatellite variation of adult frogs in the Conness, Ebbett's Pass, and Tahoe Basin populations, with corresponding plot of individual posterior probabilities for cluster membership shown at k = 3

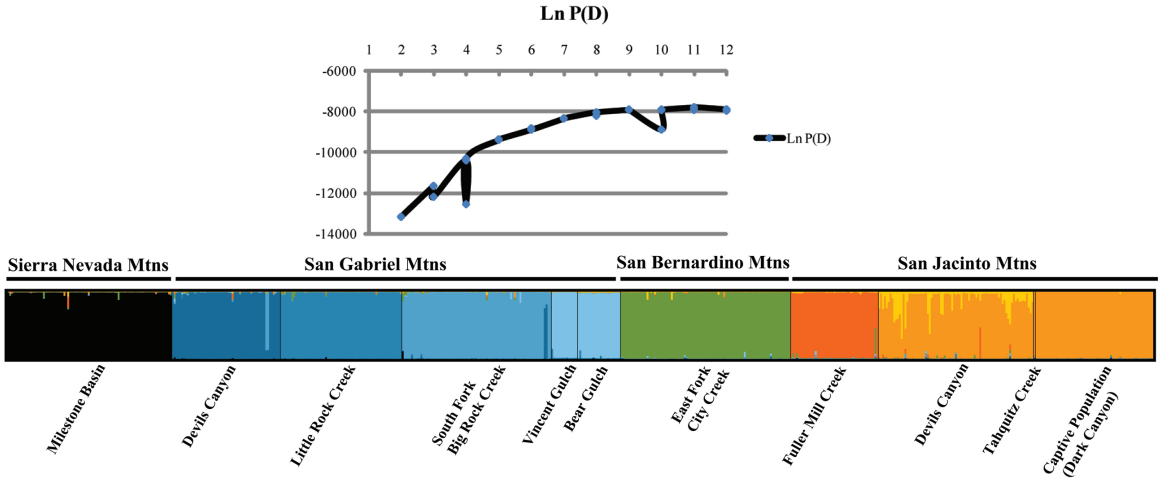


Figure 3.3: STRUCTURE analysis of microsatellite variation of adult frogs in the Milestone and Southern California populations (plot of the posterior likelihood of the data with increasing number of clusters (k), ten replicates at each k, showing a distinct plateau at k = 9), with corresponding plot of individual posterior probabilities for cluster membership shown at k = 9

	Site Name	N	AR	H_o	H_e	F
Rana sierrae	Conness	48	7.200	0.532	0.607	0.124
-	Ebbett's Pass	24	4.800	0.432	0.622	0.317†
-	Tahoe Basin	26	5.333	0.620	0.746	0.161
Rana muscosa	60 Lake Basin 1	25	4.625	0.543	0.554	0.020
-	60 Lake Basin 2	74	6.000	0.593	0.605	0.012
-	60 Lake Basin 3	25	3.625	0.447	0.451	-0.037
-	Milestone Basin	56	4.667	0.358	0.400	0.109
-	Devil's Canyon	23	3.889	0.512	0.480	-0.052
-	Little Rock Creek	40	4.111	0.434	0.521	0.178†
-	South Fork Big Creek	9	2.889	0.446	0.454	0.023
-	Vincent Gulch	12	3.333	0.511	0.426	-0.218
-	Bear Gulch	40	2.556	0.514	0.431	-0.170
-	East Fork City Creek	18	4.778	0.253	0.301	0.210†
-	Fuller-Mill Creek	29	4.111	0.651	0.577	-0.115
-	Dark Canyon	56	4.667	0.358	0.400	0.109

Table 3.1: N sample size; AR mean number of alleles; H_o observed heterozygosity; H_e expected heterozygosity; F_{is}. † indicates significance at the 0.05 level based on 10000 permutations in ARLEQUIN

	Sign test	Standardized Diff Test	Wilcoxin	M-ratio	Standard Deviation	Bd Status
Conness	0.521	0.478	0.651	0.621	0.252	+
Ebbett's Pass	0.553	0.356	0.477	0.673	0.200	−
Tahoe Basin	0.459	0.489	0.236	0.691	0.250	+
60 Lake Basin1	0.389	0.389	0.517	0.579	0.295	−
60 Lake Basin2	0.624	0.751	0.921	0.596	0.220	−
60 Lake Basin3	0.551	0.453	0.631	0.617	0.331	−
Milestone Basin	0.533	0.252	0.50	0.534	0.317	+
Milestone 1	-	-	-	0.496	0.291	−
Milestone 2	-	-	-	0.500	0.277	−
Milestone 3	-	-	-	0.459	0.292	−
Devil's Canyon	0.583	0.469	0.787	0.200	0.081	−
Little Rock Creek	0.558	0.486	0.820	0.208	0.064	−
South Fork Big Creek	0.440	0.057	0.455	0.219	0.084	−
Vincent Gulch	0.125	0.052	0.996	0.275	0.064	−
Bear Gulch	0.624	0.442	0.545	0.218	0.094	−
East Fork City Creek	0.389	**0.044**	0.150	0.174	0.112	−
Fuller-Mill Creek	0.549	0.184	0.5	0.196	0.066	−
Dark Canyon	**0.005**	0.059	**0.010**	0.179	0.037	−

Table 3.2: Bottleneck tests and Bd status in *Rana muscosa* and *Rana sierrae*. Based on simulations, M-ratio values less than 0.42 ($\theta = 1$) or 0.43 ($\theta = 10$) are significant and indicate a bottleneck. For Sign test, Standardized differences test, and Wilcoxin sign-rank test, significant differences are shown in bold. The Milestone population was also analyzed as 3 distinct populations (each corresponding to a lake) for comparison.

Name	Primer Sequence (5' to 3')	Microsatellite Motif	Length
A11_F	AACTGATACTTTTGGGTGTCTG	$(CA)_{16}$	148
A11_R	CGATTACCTGTCTTGGTGTC		
A104_F	CAACGGGGACATTCTAAAG	$(CA)_{12}$	253
A104_R	CCCCTAGTCTGCAAATAAAAA		
D11_F	GCGATACACACCCCTGAG	$(AGAT)_{11}(ATAT)(AGAT)_{17}$	196
D11_R	GAAGCGACTGGATTTTCTTG		
D14_F	TCCATGTCCATTTTGTGTTTG	$(TAGA)_{8}$	175
D14_R	GGTTACAACTGGGAGGTGTTG		
D114_F	CCTGGTGCCATTATTTTTTTAG	$(TAGA)_{21}$	236
D114_R	TTATCCCGGAGGAGTACAGTC		
D119_F	ATGCAGTTTACAGTTTCACACG	$(TAGA)_{21}$	133
D119_R	ATCCCCACACACGCTCTA		
D125_F	GGTGCTGCATCACTATAATTTC	$(TAGA)_{13}$	270
D125_R	ATGTGGACATTGGCTTTATTC		
D131_F	CCTTTGGAGGACGATACAGG	$(TAGA)_{13}$	284
D131_R	GCAGACAGTAGCACAGCACAC		
D208_F	AGTCCTTCTCCACTTTTTTCTC	$(TAGA)_{12}$	240
D208_R	CAGCCTGTTCTGGGTTATT		
D129_F	CCAAAGACAGAGGCACTTAG	$(TAGA)_{14}(TGGA)(TAGA)_{11}$	205
D129_R	TGCTCAGGACCTGTAGGTAG		

Figure 3.4: Microsatellite primers used in analysis

Chapter 4

Water flow and growth rate of the fungal pathogen *Batrachochytrium dendrobatidis*

Introduction

The fungal pathogen, *Batrachochytrium dendrobatidis* (Bd), the cause of the amphibian skin disease chytridiomycosis, has devastated amphibian populations world wide [Wake and Vredenburg, 2008, Skerratt et al., 2007, Berger et al., 1998]. Thought to be an emerging infectious disease [Skerratt et al., 2007], Bd has been observed to cause mortality in numerous frog species [Wake and Vredenburg, 2008], as well as several salamander species [Cheng et al., 2011], possibly driving several species to extinction.

In many species, Bd is highly pathogenic, with low levels of initial infection leading to high levels of mortality [Skerratt et al., 2007], while other species appear to tolerate Bd infection with little to no side effects [Weldon et al., 2004, Schloegel et al., 2010] and may serve as reservoirs for other species. California's mountain yellow-legged frog (a species complex consisting of *Rana muscosa* and *Rana sierrae*), has been particularly hard hit by the Bd epidemic, with declines of over 90% across its range [Vredenburg et al., 2007]. These highly aquatic species historically occurred in both lake and stream habitat in the high elevations (2000 meters) of the Sierra Nevada Mountains in California [Grinnell and Storer, 1924]. The introduction of non-native trout has vastly reduced and fragmented populations throughout their range [Vredenburg et al., 2007], and the introduction of Bd has all but driven the remaining populations to extinction [Briggs et al., 2005]. The earliest record of Bd infection may be in a *R. muscosa* misidentified as a *R. boylii* [Vredenburg et al., 2010].

The mechanism by which Bd kills its host remains somewhat of a mystery, but both lab experiments and field observations suggest that Bd infection impairs normal skin function [Voyles et al., 2007, 2012]. Voyles et al. [2007] found that electrolyte across the skin was impaired in infected individuals in the lab, while Voyles et al. [2012] found that infected

animals in the field had electrolyte levels associated with impaired skin function. The Bd life cycle occurs primarily in the skin of the amphibian, as Bd is believed to feed primarily on keratin or keratinized tissue [Pessier et al., 1999]. The life cycle consists of a flagellated zoospore which encysts upon the skin of the frog, forming a zoosporangia. Zoospores develop within the zoosporangia and are then released into the environment though a zoospore tube. Individual frogs may be continuously reinfected from zoospores in the environment: Rachowicz and Vredenburg [2004] found that tadpoles actually loose there infection during metamorphosis, and are reinfected as juvenile frogs, and modeling by Briggs et al. [2005] suggests that reinfection from an external zoospore pool may play an important role in driving Bd epidemics.

We conducted two studies investigating how the manipulation of the external zoospore pool via changes in flow rate affects both the growth rate of Bd on the host as well as host mortality. Our goal was to determine whether increased flow rate was associated with a decrease in growth rate of Bd on its host due to reduced reinfection, and whether this reduced reinfection rate led to lower mortality. In our first study, we simulated flow by changing animals more often than a control group. To investigate the effects of flow further, our second experiment simulated continuous flow.

Materials and Methods

Flow Experiment I

Source Population of Experimental Animals

Several clutches of *Rana muscosa* were collected from 60 Lake Basin in Kings Canyon National Park. All animals were collected as egg masses, and reared through metamorphosis at facilities at UC Berkeley.

Inoculation with B.d.

20 animals were inoculated with Bd strain LJR 119 once a day for 5 days. Animals were divided equally and housed in two 20 liter tanks during the inoculation. Inoculation was performed by flooding one Petri dish with 10 ml of water, allowing the water to stand for 10 min, and then pouring the water into the tank with the animals. The holding tank water was changed once during the inoculation period, and the animals were also fed once during this period.

Experimental Procedures

After the initial inoculation, animals were divided into two treatment groups of 10 each, one experimental and one control. Each animal was housed individually in a 4 liter mouse box (Unicage) with 200 ml of water. Water was calibrated to room temperature (15 degrees

C) by leaving a bucket in the room with the animals. Water in the experimental tanks, or "Flow" group, was changed twice per day, once in the morning and once at night, including the weekends. Animals were rinsed with fresh water before being placed in a new tank, changing gloves between each animal. Control animals were picked up and placed back in their cages each time the experimental tanks were changed. Control animals were changed once a week. All animals were fed once per week in the morning, after the flow animals had changed. Animals were fed five crickets each. All animals were swabbed once per week for the duration of the experiment, right before changing the tanks (see real time pcr below). The experiment was run for 15 weeks.

Flow Experiment II

Source Population of Experimental Animals

Because of the limited availability of experimental animals, animals that had been previously exposed to Bd were used for the second experiment. 20 R. sierrrae collected as egg masses from Ebbetts Pass (Eldorado National Forrest) and Connness Pond (Yosemite National Park). While all animals had been previously exposed to Bd, all were confirmed to be Bd negative with real time pcr before the beginning of the experiment.

Inoculation with Bd

20 animals were inoculated with B.d. strain (CJB 6) once a day for 5 days, following procedures similar to flow experiment I. All animals were divided equally and housed in two 5 gallon tanks during the inoculation. Inoculation was performed by flooding one Petri dish with 10 ml of water, allowing the water to stand for 10 min, and then pouring the water into the tank with the animals. The holding tank water was changed once during the inoculation period, and the animals were also fed once during this period.

Experimental Procedures

After the initial inoculation, animals were randomly divided into two groups of 10, one flow-though group and one recycle group. In both groups, animals were housed in 4 liter containers (Unicage) with approximately 500 ml of water, with small holes in the bottom for drainage. In the flow-through group, fresh water was continuously pumped into the top of the cage at approximately 2.5 liters per day. In the recycle group, water was pumped into the top of the cage, also at a rate of averaging 2.5 liters per day, and then collected in a container beneath the animal. Flow rates varied by individual, but rates between the two groups were similar. Flow rate for each cage was measured weekly. The collected water was then pumped back into the animal's container. The total volume of water used in the recycle group was 1.5 liters, 500 ml in the animal housing and a 1 liter reservoir. All animals were swabbed once per week, and 50ml of water was collected from the cage for filtration in order to test for zoospores in the water. During this time all tanks were cleaned a rinsed with

fresh water and the water was changed in the recycle group. The experiment was run for 10 weeks. Starting at week 2 until the end of the experiment, the flow rate of each pump was measured each week by collecting the outflow in a 60 mL flask for 5 minutes and measuring the volume of water.

Real Time PCR

Each week, every animal was swabbed to quantify the amount of Bd present on the individual by rubbing a sterile, cotton swab over the ventral surfaces and digits [Hyatt et al., 2007, Vredenburg et al., 2010]. Swabs were extracted and analyzed on a ABI StepOne Plus real time pcr machine following the protocols of Boyle et al. [2004]. In both experiments, all animals that did not show significant levels of infection (< 1 zoospore equivalent) after the first week were excluded from subsequent analysis.

Walter filtration

Each week, 50ml of water was collected from each tank and passed through a 5 micron filter using a 60 ml syringe. The filters were then dried, and then zoospore DNA was extracted using the Prepman extraction kit (Applied Biosystems), following procedures described in [Reeder et al., 2012].

Statistical Analysis

All data was analyzed in R, version 2.14.1 [R Core Team, 2012]. Kaplan-Meier estimators were used to analyze survival data using the survival library [Therneau, 2012]. The R package implements the G-rho family of Harrington and Fleming [1982], which in this case is equivalent to a log-rank test.

PCR data for both swab and filtration data was analyzed with mixed effects models using the library nlme [Pinheiro et al., 2012], using a first order auto-regressive model. Week and treatment were treated as fixed effects, with week nested within treatment. Individual was treated as a random effect with varying intercept. Since the flow rate varied across all of the cages, we also determined if there was any effect of flow on zoospore growth on the animals or zoospores found in the water. To determine whether weekly fluctuations in weekly fluctuations in flow might affect zoospore growth on the frogs in water, we looked at the one week change in flow rate compared to the change in zoospores both on the frog and in the water.

We also preformed a segmented regression using the library segmented [Muggeo, 2003, 2008] to determine if there were any breakpoints in the growth rate of Bd on the frogs, possibly due host immune response. The breakpoint model was compared with a quadratic model, and the relative goodness of fit of each model was assesed using the Akaike Information Criterion (AIC), where a decreases of > 2 indicated a better fit to the data.

Results

Flow I

Three animals in each group never showed infection levels above 0.1 zoospores, theoretically the lower bound for detection, and were removed from the analysis. Survival analysis showed significant difference in survival between the flow group and the control group (Fig 1, $p < 0.00001$). All seven animals in the control group had died by week eight, while two of seven animals from the flow group survived for the entirety of the experiment.

Animals in the control group also showed significantly higher Bd growth rates than animals in the control group (Fig 2). Though flow itself was not a significant factor in the model, the interaction of flow and time was highly significant ($p < 0.00001$). The results of the model are reported in Table 1.

Visual inspection of the residuals for the mixed effects model suggested an abrupt change in slope at week 3 of the experiment. Segmented regression analysis suggested a breakpoint at 2.5 ± 1 week (Fig 3) for the flow group, while there was no breakpoint apparent in the control group. The segmented model was significantly better fit than a simple regression model, but was not distinguishable from the model with a quadratic term.

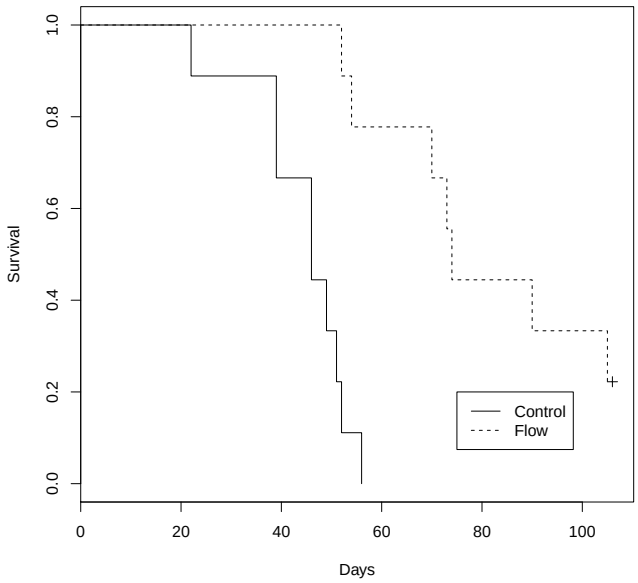
Flow II

Animals in the flow group had slightly higher survival than those in the recycle group, although the difference was not significant (Fig 3). Eighty percent of the flow group survived to the end of the experiment, as compared to fifty of the recycle group.

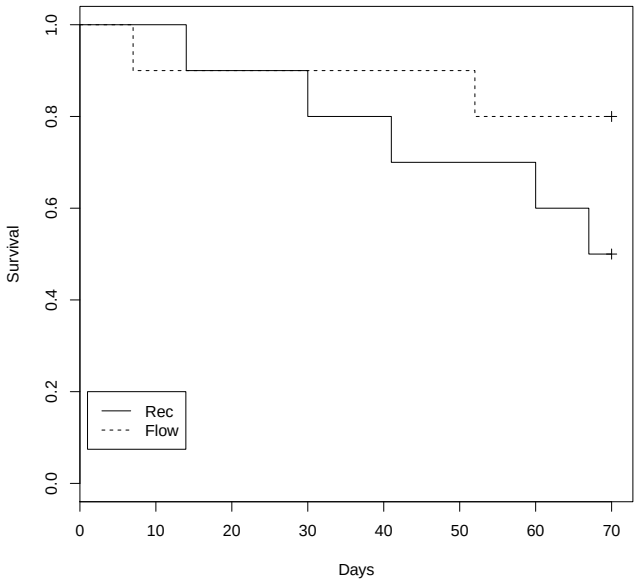
There was no detectable difference in zoospore growth rates between the two experimental groups in either the real time pcr data or the pcr based on water filtration. Animals with higher Bd loads had higher filtration values (likely because they were shedding more zoospores), although there was no difference between groups.

The average flow rate in the cages was 2.5 ± 0.9 liters per day, with a minium of .14 and a maximum of 8.1 L per day. There was no significant effect on change of flow from week to week on either zoospores measured on the skin or in the water, though there as a negative but statistically insignificant relationship between increase in flow filtered zoospores.

Overall, mortality was higher in Flow I when compared to Flow II. By week 10 in Flow I, 100% of the control group and 50% of the flow group had died, as compared to 50% of the recycle group and 20% of the flow group in the second experiment.



(a) Flow I



(b) Flow II

Figure 4.1: Kaplan-Meier survival plots for the Flow I (a) and Flow II (b) experiment2. $p < 0.00001$ for Flow I, $p = 0.2$ for Flow II. In both figures, the dotted line represents the flow through group. In figure a, the solid line represents the control group. In figure b, the solid line represent the group with water recycled.

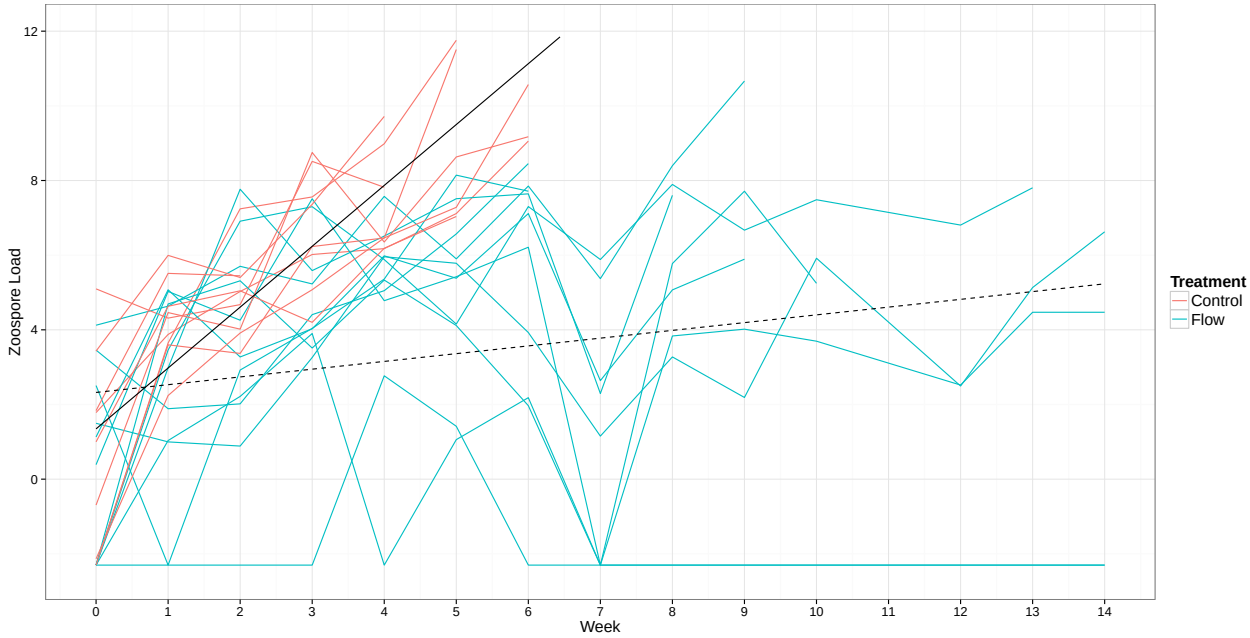


Figure 4.2: Zoospore load (log) overtime in for each individual in Flow I. Red lines are the control group, blue the control group. Solid and dash lines represent model fits for the control and flow group respectively

	Value	*se*	*p*-value
Intercept	1.3467587	0.9867918	0.1743
Week	1.6303872	0.2166024	0.00001
Flow Treatment	0.9757719	1.2781617	0.4551
Treatment:Week	-1.4225823	0.2296019	0.00001

Table 4.1: Results from mixed effects model for Flow I: effect size, standard error, and *p*-values

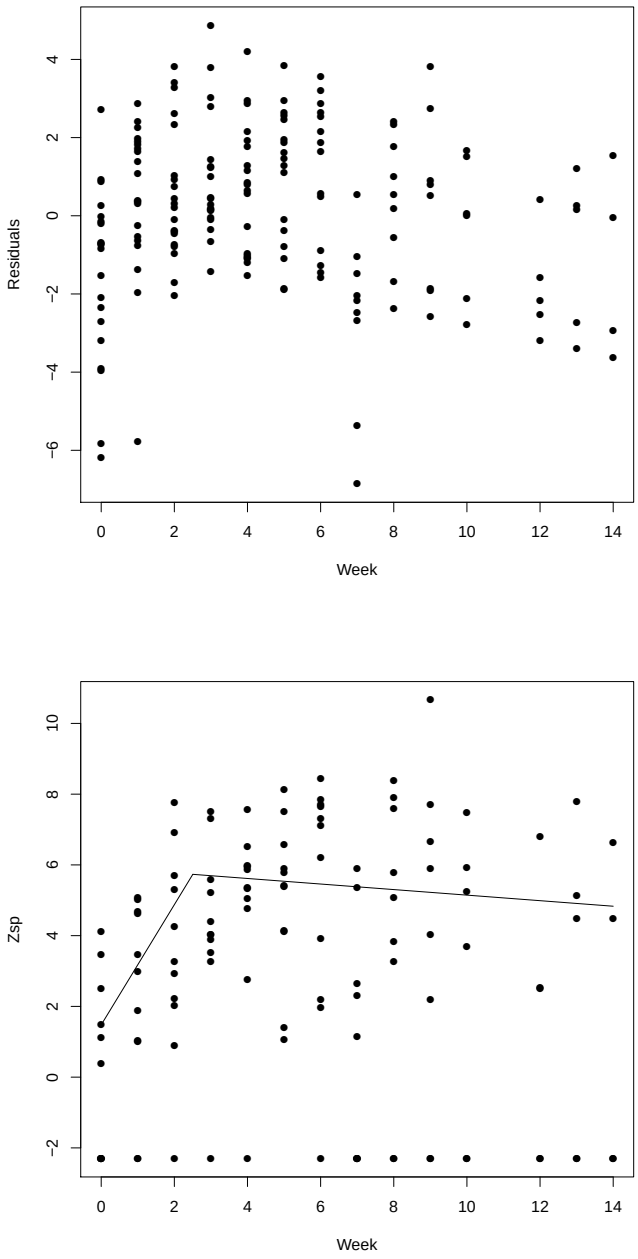


Figure 4.3: Top: residual plot of the mixed effects model, indicating a break point around week 3. Bottom: Zoospore load (log(x+0.1)) over time in for the flow group in Flow I, with estimated breakpoint at 2.5 ± 1 weeks

Discussion

We found that manipulation of the flow rate in naive frogs exposed to Bd significantly reduced the growth rate of Bd as well as the time to mortality of the frogs(Flow I, Fig 2). However, there was no significant effect when the flow rate was manipulated in frogs previously exposed to Bd (Flow II, Fig 2). One possible explanation is that completely removing the water in the first experiment, while representing a lower daily turnover in total water flow, actually completely purged the zoospores from the container, while the flow rate in experiment 2 was insufficient to reduce the zoospore pool. Bd zoospores are flagellated, and actively swim in the aquatic environment [Piotrowski et al., 2004], so it is possible that certain flow rates may not affect zoospore density, especially at the low rates simulated in Flow exp II. This is supported by the fact that in Flow II we found no effect of flow rate on zoospore density in the cage water. It is possible that the flow rates simulated in this experiment were insufficient to reduce the external zoospore density.

It is certainly possible that the previous exposure of the animals in Flow Exp II resulted in lower Bd growth rate overall, and thus lower mortality. Although we were unable to distinguish between the breakpoint model and the quadratic model, we found a decline in growth rate after 2.5 weeks (Fig 3) in Flow I. This could be consistent with an immune response after the initial exposure to Bd. Animals in Flow II had a lower motality overall than animals in Flow I, even within the control groups, which should have had similar exposure to Bd. Since all animals in Flow II had previously been exposed, it is possible that aquired immunity led to a reduction in mortality in Flow II.

The results of these experiments suggest that flow rate can reduce the growth rate of Bd on its amphibian host in certain circumstances. Because of naive animals were unavailable for our second experiment, we were unable to determinte how continous flow rate affects growth rate of Bd. However our results indicate that flows greater than those simulated in this experiment would be necessary to reduce growth rate of Bd on its host.

Bibliography

R.M. Anderson and R.M. May. *Infectious Diseases of Humans: Dynamics and Control*. Science Publications, Oxford, 1991.

A.R. Backlin, C.J. Hitchcock, R.N. Fisher, P. Warburton, M.L.and Trenham, S.A. Hathaway, and C.S. Brehme. Natural history and recovery analysis for southern california populations of the mountain yellow-legged frog (rana muscosa), 2003. *US Geological Survey Final Report prepared for the California Department of Fish and Game, Angeles and San Bernardino National Forests*, 2004.

Frank Ball. Deterministic and stochastic epidemics with several kinds of susceptibles. *Advances in Applied Probability*, 17:1–22, 1985.

Lee Berger, Rick Speare, Peter Daszak, D. Earl Green, Andrew A. Cunningham, C. Louise Goggin, Ron Slocombe, Mark A. Ragan, Alex D. Hyatt, Keith R. McDonald, Harry B. Hines, Karen R. Lips, Gerry Marantelli, and Helen Parkes. Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of australia and central america. *Proceedings of the National Academy of Sciences*, 95(15):9031–9036, 1998.

Robert Edmund Bingham. *Differentiation Across Multiple Spatial Scales in Three California Amphibians*. PhD thesis, University of California Berkeley, 2007.

D. G. Boyle, D. B. Boyle, V. Olsen, J. A. T. Morgan, and A. D. Hyatt. Rapid quantitative detection of chytridiomycosis (batrachochytrium dendrobatidis) in amphibian samples using real-time taqman pcr assay. *Diseases of Aquatic Organisms*, 60(2):141–148, August 9 2004.

Cheryl J. Briggs, Vance T. Vredenburg, Roland A. Knapp, and Lara J. Rachowicz. Investigating the population-level effects of chytridiomycosis: An emerging infectious disease of amphibians. *Ecology (Washington D C)*, 86(12):3149–3159, Dec 2005.

Cheryl J. Briggs, Roland A. Knapp, and Vance T. Vredenburg. Enzootic and epizootic dynamics of the chytrid fungal pathogen of amphibians. *Proceedings of the National Academy of Sciences*, 107(21):9695–9700, 2010. doi: 10.1073/pnas.0912886107. URL `http://www.pnas.org/content/107/21/9695.abstract`.

Tina L. Cheng, Sean M. Rovito, David B. Wake, and Vance T. Vredenburg. Coincident mass extirpation of neotropical amphibians with the emergence of the infectious fungal pathogen batrachochytrium dendrobatidis. *Proceedings of the National Academy of Sciences*, 108 (23):9502–9507, 2011. doi: 10.1073/pnas.1105538108.

Louns Chikhi, Vitor C. Sousa, Pierre Luisi, Benoit Goossens, and Mark A. Beaumont. The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics*, 186(3):983–995, November 2010. doi: 10.1534/genetics.110.118661.

J. M. Cornuet and G. Luikart. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144(4):2001–2014, DEC 1996. PT: J.

Hans D. Daetwyler, Beatriz Villanueva, and John A. Woolliams. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS ONE*, 3(10):e3395, 10 2008. doi: 10.1371/journal.pone.0003395.

O. Diekmann and J. A. P. Heesterbeek. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation (Wiley Series in Mathematical & Computational Biology)*. Wiley, 1 edition, May 2000. ISBN 0471492418. URL `http://www.worldcat.org/isbn/0471492418`.

O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz. On the definition and the computation of the basic reproduction ratio r_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28:365–382, 1990. ISSN 0303-6812. URL `http://dx.doi.org/10.1007/BF00178324`. 10.1007/BF00178324.

Andrea B. Doeschl-Wilson, R. Davidson, J. Conington, T. Roughsedge, M. R. Hutchings, and B. Villanueva. Implications of host genetic variation on the risk and prevalence of infectious diseases transmitted through the environment. *Genetics*, 188(3):683–693, 2011. doi: 10.1534/genetics.110.125625. URL `http://www.genetics.org/content/188/3/683.abstract`.

W.J. Ewens. *Mathematical population genetics*. Springer, 2004.

Gentile Francesco Ficetola, Trenton W. J. Garner, and Fiorenza De Bernardi. Genetic diversity, but not hatching success, is jointly affected by postglacial colonization and isolation in the threatened frog, rana latastei. *Molecular Ecology*, 16(9):1787–1797, 2007. ISSN 1365-294X. doi: 10.1111/j.1365-294X.2006.03198.x. URL `http://dx.doi.org/10.1111/j.1365-294X.2006.03198.x`.

W. Chris Funk, Michael S. Blouin, Paul Stephen Corn, Bryce A. Maxell, David S. Pilliod, Stephen Amish, and Fred W. Allendorf. Population structure of columbia spotted frogs (rana luteiventris) is strongly affected by the landscape. *Molecular Ecology*, 14(2):483–496,

2005. ISSN 1365-294X. doi: 10.1111/j.1365-294X.2005.02426.x. URL `http://dx.doi.org/10.1111/j.1365-294X.2005.02426.x`.

J. C. Garza and E. G. Williamson. Detection of reduction in population size using data from microsatellite loci. *Molecular ecology*, 10(2):305–318, FEB 2001. PT: J.

W.M. Getz and J. Pickering. Epidemic models: thresholds and population regulation. *American Naturalist*, 121:892–898, 1983.

J. Grinnell and T. Storer. *Animal life in the Yosemite*. University of California Press, Berkeley, 1924.

David P. Harrington and Thomas R. Fleming. A class of rank test procedures for censored survival data. *Biometrika*, 69(3):553–566, 1982. doi: 10.1093/biomet/69.3.553.

Joel N. Hirschhorn and Mark J. Daly. Genome-wide association studies for common diseases and complex traits. *Nature reviews. Genetics*, 6(2):95–108, print 2005. URL `http://dx.doi.org/10.1038/nrg1521`. M3: 10.1038/nrg1521; 10.1038/nrg1521.

Peter Hraber, Carla Kuiken, and Karina Yusim. Evidence for human leukocyte antigen heterozygote advantage against hepatitis c virus infection. *Hepatology*, 46(6):1713–1721, 2007. ISSN 1527-3350. doi: 10.1002/hep.21889. URL `http://dx.doi.org/10.1002/hep.21889`.

A. D. Hyatt, D. G. Boyle, V. Olsen, D. B. Boyle, L. Berger, D. Obendorf, A. Dalton, K. Kriger, M. Hero, H. Hines, R. Phillott, R. Campbell, G. Marantelli, F. Gleason, and A. Colling. Diagnostic assays and sampling protocols for the detection of batrachochytrium dendrobatidis. *Diseases of Aquatic Organisms*, 73(3):175–192, Jan 18 2007.

Pauline Kamath and Wayne Getz. Adaptive molecular evolution of the major histocompatibility complex genes, dra and dqa, in the genus equus. *BMC Evolutionary Biology*, 11(1): 128, 2011.

Curtis M. Lively. The effect of host genetic diversity on disease spread. *The American Naturalist*, 175(6):pp. E149–E152, 2010. ISSN 00030147. URL `http://www.jstor.org/stable/10.1086/652430`.

J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438:355–359, 2006.

Kirk E. Lohmueller, Celeste L. Pearce, Malcolm Pike, Eric S. Lander, and Joel N. Hirschhorn. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature genetics*, 33(2):177–182, print 2003. URL `http://dx.doi.org/10.1038/ng1071`. M3: 10.1038/ng1071; 10.1038/ng1071.

E. Luquet, P. David, j.-P. Lena, P. Joly, L. Konecny, C. DUFRESNES, N. Perrin, and S. Plenet. Heterozygosity??fitness correlations among wild populations of european tree frog (hyla arborea) detect fixation load. *Molecular Ecology*, 20(9):1877–1887, 2011. ISSN 1365-294X.

Emilien Luquet, Trenton W.J. Garner, Jean-Paul Lna, Christophe Bruel, Pierre Joly, Thierry Lengagne, Odile Grolet, and Sandrine Plnet. Genetic erosion in wild populations makes resistance to a pathogen more costly. *Evolution*, 66(6):1942–1952, 2012. ISSN 1558-5646. doi: 10.1111/j.1558-5646.2011.01570.x. URL `http://dx.doi.org/10.1111/j.1558-5646.2011.01570.x`.

KirstenJ. Monsen and MichaelS. Blouin. Extreme isolation by distance in a montane frog rana cascadae. *Conservation Genetics*, 5:827–835, 2004. ISSN 1566-0621. doi: 10.1007/s10592-004-1981-z. URL `http://dx.doi.org/10.1007/s10592-004-1981-z`.

Vito M.R. Muggeo. Estimating regression models with unknown break-points. *Statistics in Medicine*, 22:3055–3071, 2003.

Vito M.R. Muggeo. segmented: an r package to fit regression models with broken-line relationships. *R News*, 8(1):20–25, 2008. URL `http://cran.r-project.org/doc/Rnews/`.

Allan P. Pessier, Donald K. Nichols, Joyce E. Longcore, and Melvin S. Fuller. Cutaneous chytridiomycosis in poison dart frogs (dendrobates spp.) and white's tree frogs (litoria caerulea). *Journal of Veterinary Diagnostic Investigation*, 11(2):194–199, 1999. doi: 10.1177/104063879901100219. URL `http://vdi.sagepub.com/content/11/2/194.short`.

Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, and R Core Team. *nlme: Linear and Nonlinear Mixed Effects Models*, 2012. R package version 3.1-104.

Jeffrey S. Piotrowski, Seanna L. Annis, and Joyce E. Longcore. Physiology of batrachochytrium dendrobatidis, a chytrid pathogen of amphibians. *Mycologia*, 96(1):9–15, 2004. URL `http://www.mycologia.org/content/96/1/9.abstract`.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL `http://www.R-project.org/`. ISBN 3-900051-07-0.

L. Rachowicz and Vance T. Vredenburg. Transmission of batrachochytrium dendrobatidis within and between amphibian life stages. *Diseases of Aquatic Organisms*, 61:75–83, 2004.

Natalie M. M. Reeder, Allan P. Pessier, and Vance T. Vredenburg. A reservoir species for the emerging amphibian pathogen ¡italic¿batrachochytrium dendrobatidis¡/italic¿ thrives in a landscape decimated by disease. *PLoS ONE*, 7(3):e33567, 03 2012. doi: 10.1371/journal.pone.0033567.

L. M. Schloegel, C. M. Ferreira, T. Y. James, M. Hipolito, J. E. Longcore, A. D. Hyatt, M. Yabsley, A. M. C. R. P. F. Martins, R. Mazzoni, A. J. Davies, and P. Daszak. The north american bullfrog as a reservoir for the spread of batrachochytrium dendrobatidis in brazil. *Animal Conservation*, 13:53–61, 2010. ISSN 1469-1795. doi: 10.1111/j.1469-1795.2009.00307.x. URL `http://dx.doi.org/10.1111/j.1469-1795.2009.00307.x`.

Z. Shen, F. Ning, W. Zhou, X. He, C. Lin, D. P. Chin, Z. Zhu, and A Schuchat. Super-spreading sars events, beijing, 2003. *Emerg Inf Dis*, 10:256–260, 2004.

Lee Skerratt, Lee Berger, Richard Speare, Scott Cashins, Keith McDonald, Andrea Phillott, Harry Hines, and Nicole Kenyon. Spread of chytridiomycosis has caused the rapid global decline and extinction of frogs. *EcoHealth*, 4(2):125–134, 06/01 2007. M3: 10.1007/s10393-007-0093-5.

R. W. Slade and H. I. McCallum. Overdominant vs. frequency-dependent selection at mhc loci. *Genetics*, 132(3):861–862, 1992.

A. J. Springbett, K. MacKenzie, J. A. Woolliams, and S. C. Bishop. The contribution of genetic diversity to the spread of infectious diseases in livestock populations. *Genetics*, 165 (3):1465–1474, 2003. URL `http://www.genetics.org/content/165/3/1465.abstract`.

N. Takahata and M. Nei. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics*, 124:967–978, 1990.

Terry Therneau. *A Package for Survival Analysis in S*, 2012. R package version 2.36-12.

J. Voyles, L. Berger, S. Young, R. Speare, R. Webb, J. Warner, D. Rudd, R. Campbell, and L. F. Skerratt. Electrolyte depletion and osmotic imbalance in amphibians with chytridiomycosis. *Dis Aquat Org*, 77(2):113–118, 2007. URL `http://www.int-res.com/abstracts/dao/v77/n2/p113-118/`. 10.3354/dao01838.

Jamie Voyles, Vance T. Vredenburg, Tate S. Tunstall, John M. Parker, Cheryl J. Briggs, and Erica Bree Rosenblum. Pathophysiology in mountain yellow-legged frogs (*Rana muscosa*) during a chytridiomycosis outbreak. *PLoS ONE*, 7(4):e35374, 04 2012. doi: 10.1371/journal.pone.0035374. URL `http://dx.doi.org/10.1371%2Fjournal.pone.0035374`.

V. T. Vredenburg, R. Bingham, R. Knapp, J. A. T. Morgan, C. Moritz, and D. Wake. Concordant molecular and phenotypic data delineate new taxonomy and conservation priorities for the endangered mountain yellow-legged frog. *Journal of Zoology (London)*, 271(4):361–374, Apr 2007.

Vance T. Vredenburg, Roland A. Knapp, Tate S. Tunstall, and Cheryl J. Briggs. Dynamics of an emerging disease drive large-scale amphibian population extinctions. *Proceedings of the National Academy of Sciences*, 107(21):9689–9694, May 25 2010.

David B. Wake and Vance T. Vredenburg. Are we in the midst of the sixth mass extinction? a view from the world of amphibians. *Proceedings of the National Academy of Sciences*, 105(Supplement 1):11466–11473, 2008. doi: 10.1073/pnas.0801921105.

K. Mathias Wegner, Martin Kalbe, Helmut Schaschl, and Thorsten B.H. Reusch. Parasites and individual major histocompatibility complex diversity-an optimal choice? *Microbes and Infection*, 6(12):1110 – 1116, 2004. ISSN 1286-4579. doi: 10.1016/j.micinf.2004.05.025. URL `http://www.sciencedirect.com/science/article/pii/S1286457904002187`.

Che Weldon, Louis H. du Preez, Alex D. Hyatt, Reinhold Muller, and Rick Speare. Origin of the amphibian chytrid fungus. *Emerging Infectious Diseases*, 10(12):2100–2105, December 2004.

Hadley Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009. ISBN 978-0-387-98140-6. URL `http://had.co.nz/ggplot2/book`.

EllenG. Williamson-Natesan. Comparison of methods for detecting bottlenecks from microsatellite loci. *Conservation Genetics*, 6:551–562, 2005. ISSN 1566-0621. doi: 10.1007/s10592-005-9009-5. URL `http://dx.doi.org/10.1007/s10592-005-9009-5`.

G. A. Wilson, T. L. Fulton, K. Kendell, G. Scrimgeour, C. A. Paszkowski, and D. W. Coltman. Genetic diversity and structure in canadian northern leopard frog (rana pipiens) populations: implications for reintroduction programs. *Canadian Journal of Zoology*, 86 (8):863–874, 2008. doi: 10.1139/Z08-062.

Andrew Yates, Rustom Antia, and Roland R Regoes. How do pathogen evolution and host heterogeneity interact in disease emergence? *Proceedings of the Royal Society B: Biological Sciences*, 273(1605):3075–3083, 2006. doi: 10.1098/rspb.2006.3681.

Aibin Zhan, Cheng Li, and Jinzhong Fu. Big mountains but small barriers: Population genetic structure of the chinese wood frog (rana chensinensis) in the tsinling and daba mountain region of northern china. *BMC Genetics*, 10(1):17, 2009. ISSN 1471-2156. doi: 10.1186/1471-2156-10-17. URL `http://www.biomedcentral.com/1471-2156/10/17`.

S. Zhao, Q. Dai, and J. Fu. Do rivers function as genetic barriers for the plateau wood frog at high elevations? *Journal of Zoology*, 279(3):270–276, 2009. ISSN 1469-7998. doi: 10.1111/j.1469-7998.2009.00615.x. URL `http://dx.doi.org/10.1111/j.1469-7998.2009.00615.x`.