AN ABSTRACT OF THE DISSERTATION OF

Ivan C. Phillipsen for the degree of Doctor of Philosophy in Zoology presented on June 2, 2010.

Title: Population Genetics of Ranid Frogs: Investigating Effective Population Size and Gene Flow

Abstract approved:

Michael S. Blouin

This dissertation focuses on the evolutionary forces of genetic drift and gene flow in frog populations. The balance of these two forces and the force of mutation largely determine the amount of neutral genetic variation within populations as well as the degree of genetic similarity among populations. The stochastic evolutionary change caused by genetic drift can be quantified through the use of the effective population size (N_e) parameter. The effective size of a population is the number of breeding individuals in a conceptual, ideal population that would evolve by genetic drift at the same rate as the real population being studied. How a population responds to mutation, selection, and gene flow depends on N_e, rather than the actual census population size (N). In most natural populations, N_e is considerably smaller than N. For these reasons, N_e is a fundamental parameter in basic population genetics theory as well as in applied conservation genetics. The degree of neutral genetic similarity between populations is highly dependent upon gene flow. When gene flow between a pair of populations is low, the populations are likely to become genetically

differentiated. Conversely, when gene flow between populations is high, the populations will tend to be more genetically similar.

Amphibians are good model organisms for studying genetic drift and gene flow because they tend to exhibit strong population structure at small spatial scales. This is a consequence of their generally small population sizes, natal philopatry, limited dispersal capabilities, and restricted habitat requirements. They are expected to have easily-detectable signatures of spatial genetic structure and genetic drift. Amphibians can be used as models to further our understanding of evolutionary processes and that understanding can be applied to the conservation of amphibians. Equipped with knowledge of what naturally influences genetic drift and gene flow in amphibians, we can apply the principles of population genetics to mitigate the genetic consequences of amphibian declines.

In Chapters 2 and 3, I used molecular genetic data from frog populations to investigate N_e and the related parameter N_b (the effective number of breeders). Chapter 2 is a study of a single population of the Oregon spotted frog (*Rana pretiosa*). My aim was to determine where in the life cycle of this species the greatest reduction in N_b occurs. I used genetic data from microsatellites to estimate N_b at two different life stages, eggs and metamorphs, and found that estimates of N_b were similar at both stages. This result suggests that inflated variance in family size due to egg mass mortality is not a primary cause of N_e reductions relative to N in this species. Chapter 3 is a comparison of N_e estimates within and among four species of frogs in the family Ranidae: *R. pretiosa*, *R. luteiventris*, *R. cascadae*, and *Lithobates pipiens*. I obtained

N_e estimates for 90 populations across the four species, using microsatellite data and several different estimators. The first three species and the western populations of *L. pipiens* have very small effective sizes (< 50). Eastern populations of *L. pipiens* are much larger, with N_e estimates in the hundreds and thousands. I also found significant correlations between N_e estimates and latitude, longitude, or altitude in *R. luteiventris* and *L. pipiens*.

Chapter 4 is a study of gene flow among populations of the Cascades frog (*Rana cascadae*) in the Olympic Mountains of Washington. I quantified genetic differentiation among 22 *R. cascadae* populations with data from microsatellite markers and used a landscape genetics approach to identify environmental features that have strong influences on gene flow in this species. I used a Random Forests statistical procedure to assess which of several structural connectivity models and 15 landscape variables explained the most variation in genetic distances among populations. I found that the best-fitting Random Forests models were based on different structural connectivity models for two datasets: 'within' and 'between' genetic clusters of populations. The landscape variables identified as the most important also differed across the two datasets, suggesting that landscape influences vary across spatial scales.

The results presented in this dissertation led to an increased understanding of effective population size in ranid frogs and of the environmental factors that influence population structure in *R. cascadae*. These studies provide a foundation for further

research on the specific factors that influence genetic drift and gene flow in these species.

Population Genetics of Ranid Frogs: Investigating Effective Population Size and Gene Flow

by
Ivan C. Phillipsen

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 2, 2010
Commencement June, 2011

Doctor of Philosophy dissertation of Ivan C. Phillipsen presented on June 2, 2010.

APPROVED:

Major Professor, representing Zoology

Chair of the Department of Zoology

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Ivan C. Phillipsen, Author

ACKNOWLEDGEMENTS

My years as a PhD student at Oregon State University have been the most exhilarating and happiest of my life. I feel extremely privileged to have had this opportunity to pursue my dream of being an evolutionary biologist. I owe my thanks to the scores of charming frogs I encountered through the course of my research (sorry about the toes!) as well as a bunch of humans who have made this all possible and enjoyable.

First, I'd like to thank my advisor Mike Blouin. As someone who wanted to study amphibian population genetics, I was lucky to be welcomed into Mike's lab, which now focuses on salmon and parasite study systems instead of frogs (as it had in years past). From the beginning of my time here, Mike was enthusiastic about my research ideas and always seemed happy to dive back into the world of frog genetics with me. He has been a great support, providing indispensible guidance in my research as well as advice on how to navigate the world of academia. I appreciate Mike's patience with me through the vicissitudes of my journey to becoming a more efficient and independent researcher. I thank Mike for providing me with all the resources I needed to pursue my interests and carry out my various research projects. I look forward to continuing our professional relationship as well as our friendship.

I would like to thank to my graduate committee. My research has benefitted greatly from the input of Steve Arnold, Tiff Garcia, Julia Jones, and Dave Lytle. I appreciate all the helpful suggestions they gave me at key junctures in my graduate

career. Each of them has been exceptionally kind and supportive and I feel fortunate to have had the guidance of such a ‘supergroup’ of scientists.

All of the Blouin Lab folk have made day to day work life a real pleasure. Jacob Tennessen and Mark Christie were awesome office mates and friends. I enjoyed brainstorming research ideas with them as well as laughing about trivial nonsense. Kaitlin Bonner has been a great, supportive friend who brought a lot of positive energy to the lab. Becky Cooper, Vindhya Amarasinghe, and Melanie Marine provided critical help with the logistics of my labwork. Hitoshi Araki was an inspiration as a scientist and incredibly nice guy.

I thank Joe Beatty and all of the Zoology Office staff—Tara Bevandich, Traci Durrell-Khalife, and Tori Givigliano— for their dedication to the graduate students. On countless occasions they helped me with practical matters and smoothed the path for me as I negotiated the paperwork and hoops of graduate student life.

The Zoology graduate students and postdocs are an amazing group of people and I am lucky to count so many friends among them. I have shared many outstanding, happy moments with them. I am grateful for the emotional and intellectual support they have given me over the years. For their friendship and camaraderie, I especially want to thank Betsy Bancroft, Kate Boersma, Mike Bogan, Angela Brandt, Sarah Eddy, Dafne Eerkes-Medrano, Deb Finn, Chris Friesen, Steph Gervasi, Margo Hessing-Lewis, Karen Kiemnec-Tyburczy, Laura McMullen, Lindsay Michael-Biga, Santiago Perez, Christy Schnitzler, Catherine Searle, Joe Tyburczy, Josef Uyeda,

Elyse Vaccaro, Mike Westphal, Elisha Wood-Charleson, Asako Yamamuro, and Phoebe Zarnetske.

Outside of the Zoology Department, there are a few denizens of Corvallis and beyond who cannot go unrecognized. I wholeheartedly thank Jay Zarnetske, Mike Gallagher, Joshua Weber, Jeremy Noring, Sam Hartsfield, and Cassondra Pittman for the friendship and love they have given me.

Lastly and most importantly, I want to thank my parents, Gregg and Katie, for their limitless, encouragement, and love.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF APPENDIX FIGURES

DEDICATION

This is for my Mom and Dad

Population Genetics of Ranid Frogs: Investigating Effective Population Size and Gene Flow

CHAPTER 1: GENERAL INTRODUCTION

The incredible diversity of life on Earth is the ultimate outcome of processes that generate, maintain, and distribute genetic variation at the level of populations within species. The amount and forms of genetic variation harbored by a population are the products of evolutionary forces acting over time. There is also a spatial dimension to the actions of these forces— few natural populations are completely isolated, such that genetic variation in a species is typically distributed among multiple populations in a landscape. Signatures of many evolutionary processes are carried in the genomes of individuals in a population. These can be detected and deciphered using modern molecular genetic methods. This dissertation focuses on the evolutionary forces of genetic drift and gene flow in frog populations. The balance of these two forces largely determines the amount of genetic variation within populations as well as the degree of genetic similarity among populations (Wright, 1931).

Genetic drift is the random change in allele frequencies that occurs from one generation to the next because alleles in a generation of offspring are only a random, imperfect sample of the alleles in the parental generation (Wright, 1931). Thus, evolutionary change caused by genetic drift is stochastic. Genetic drift is strongest in small populations, where the effect of sampling error between generations is most extreme. Genetic drift must be quantified in order to infer its role in the history of a population or to predict its effect on the continuing evolution of a population. The

parameter used to quantify genetic drift is the effective population size (N_e). The effective size of a population is the number of breeding individuals in a conceptual, ideal population that would evolve by genetic drift at the same rate as the real population being studied (Wright, 1931; Charlesworth 2009). The strength of genetic drift is inversely proportional to N_e, i.e. drift is strong when N_e is small and vice versa. How a population responds to mutation, selection, and gene flow depends on N_e, rather than the actual census population size (N). In most natural populations, N_e is considerably smaller than N (Frankham, 1995). For these reasons, N_e is a fundamental parameter in basic population genetics theory as well as in applied conservation genetics.

Most species are represented by multiple populations that differ in their degrees of genetic similarity. The degree of similarity between populations is highly dependent upon the magnitude of gene flow. When few individuals are exchanged via migration between a pair of populations, such that gene flow is low, the populations are likely to become genetically differentiated. Conversely, when gene flow between populations is high, the populations evolve more like a single genetic unit.

Amphibians are good model organisms for studying genetic drift and gene flow because they tend to exhibit high levels of genetic differentiation among populations at small geographic scales (Beebee, 2005; Chan and Zamudio, 2009). This is a consequence of their generally small population sizes, natal philopatry, limited dispersal capabilities, and restricted habitat requirements (Waldman and McKinnon, 1993). Compared to other terrestrial vertebrates, amphibians are expected to have

easily-detectable signatures of spatial genetic structure and genetic drift. Pond-breeding species have relatively discrete habitats, which facilitates the modeling of the spatial structure of their populations. In terms of practicality, amphibians are often easy to locate, handle, and sample in the field.

Amphibians can be used as models to further our understanding of evolutionary processes and that understanding can, in turn, be applied to the conservation of amphibians. A majority of amphibian species around the world face an increasing risk of extinction, due to a combination of threats such as habitat destruction, pollution, disease, and invasive species (Stuart et al, 2004). Small, declining populations inhabiting fragmented habitats are likely to suffer losses of genetic diversity through drift and reduced gene flow. Equipped with knowledge of what naturally influences genetic drift and gene flow in amphibians, we can apply the principles of population genetics to mitigate the genetic consequences of amphibian declines.

In Chapters 2 and 3, I used molecular genetic data from frog populations to investigate N_e and the related parameter N_b (the effective number of breeders), which applies to only the breeding adults of a population in a single reproductive season. Chapter 2 is a study of a single population of the Oregon spotted frog (*Rana pretiosa*), a declining amphibian in the Pacific Northwest. My aim was to determine where in the life cycle of this species the greatest reduction in N_b occurs. Several demographic factors (e.g. unequal sex ratio, fluctuating population size, and nonrandom variance in family size) can reduce N_e (or N_b) relative to N. I used genetic

data from microsatellites to estimate N_b at two different life stages: eggs and metamorphs. Knowing which demographic factors have the greatest influence in populations of an imperiled species can provide insight into the best management solutions for maximizing N_e, thereby minimizing the loss of genetic diversity through drift. I found that estimates of N_b were similar at both life stages, suggesting that inflated variance in family size due to egg mass mortality may not be a primary cause of N_e reductions relative to N in this species.

Chapter 3 is a comparison of N_e estimates within and among four species of frogs in the family Ranidae: *R. pretiosa, R. luteiventris, R. cascadae,* and *Lithobates pipiens*. Only recently, with the increasing accessibility of molecular genetic data and the development of powerful analytical methods, has it become relatively easy to estimate N_e for many populations. I obtained N_e estimates for 90 populations across the four species, using microsatellite data and several different estimators. My objectives were to: (1) determine the typical N_e estimate values for each species, (2) determine the strength of the correlation between genetic diversity and N_e estimates, (3) compare N_e estimates among the species and offer hypotheses to explain the differences, and (4) test for correlations between each of several geographic variables and N_e estimates within each species. The first three species and the westernmost populations of *L. pipiens* have very small effective sizes, less than 50. Eastern populations of *L. pipiens* are much larger, with N_e estimates in the hundreds and thousands. I also found significant correlations between N_e estimates and latitude, longitude, or altitude in *R. luteiventris* and *L. pipiens*.

Chapter 4 is a study of gene flow among populations of the Cascades frog (*Rana cascadae*) in the Olympic Mountains of Washington. I quantified genetic differentiation among 22 *R. cascadae* populations with data from microsatellite markers and used a landscape genetics approach to identify environmental features that have strong influences on gene flow in this species. I constructed three alternative models of connectivity among populations of *R. cascadae* in the Olympic Mountains: one based on linear (i.e. Euclidean) connections, one based on a minimum spanning tree network of pond habitats, and one based on the connectivity of stream drainages. I used a Random Forests statistical procedure to assess which of these models explained the most variation in genetic distances among populations. For each structural connectivity model, 15 Landscape variables were measured along paths linking pairs of populations. I wanted to identify which of these variables, given a particular connectivity model, are the most important predictors of genetic differentiation in *R. cascadae*. I evaluated these associations both within and between genetic clusters of populations, in order to determine how the importance of landscape variables differs with spatial scale. I found that the best-fitting Random Forests models were based on different structural connectivity models for the within and between group datasets. The landscape variables identified as the most important also differed across the two datasets, suggesting that landscape influences vary across spatial scales.

CHAPTER 2

Effective number of breeding adults in Oregon spotted frogs (*Rana pretiosa*): genetic estimates at two life stages

Ivan C. Phillipsen, Jay Bowerman, and Michael Blouin

Conservation Genetics

In Press: DOI 10.1007/s10592-009-9862-8

Abstract

We used genetic methods to estimate the effective number of breeders (N_b) in a population of *Rana pretiosa*, an imperiled amphibian in western North America. Microsatellite data was gathered from large samples of adults, eggs, and juveniles collected in 2006. We wished to determine where in the life cycle the greatest reductions in N_b occur, and to compare genetic estimates of N_b to an egg mass count estimate of the number of breeding adults. We predicted that N_b estimated at the metamorph stage would be reduced by increased variance in family size due to egg mass mortality. Contrary to our prediction, estimates of N_b at the egg and metamorph stages were similar. Thus, we found no evidence of inflated variance in family size between the two stages. If our results for this population are typical for *R. pretiosa*, then increased variance in family size during the egg to metamorph stage may not be a strong factor in reducing the effective population sizes (N_e) relative to the census sizes (N) in this species.

Introduction

Effective population size (N_e) is a fundamental parameter in the theory and practice of conservation genetics. Related to N_e is the effective number of breeders, N_b, a parameter influenced by most of the same demographic factors as N_e but which applies to only the breeding adults of a population in a single reproductive season. Estimates of N_b or N_e in natural populations are usually much lower than the census population size, N (e.g. Frankham, 1995). What causes N_e and N_b to be lower than N is not well understood for many species.

The ongoing loss of global amphibian diversity is a widely recognized ecological crisis (Stuart *et al*. 2004). Values of N_e/N and N_b/N reported for amphibians range widely, from 0.001 (Easteal 1985) to greater than 0.7 (Brede and Beebee 2006). What features of the life histories of different species might predispose them to have different ratios? For example, there is some intriguing evidence that toads of the genus *Bufo* have N_e/N ratios an order of magnitude lower than those of frogs of the genus *Rana* (Hoffman *et al*. 2004; Brede and Beebee 2006). Understanding what factors in the life cycle of amphibians are most responsible for reductions in N_b or N_e could be very useful for managing loss of genetic diversity in these taxa.

The two factors thought to most dramatically reduce N_e in animal populations are fluctuating population size and non-random variance in family size (Frankham 1995). Pond-breeding frogs may be particularly susceptible to reductions in N_e by these factors. Populations of frogs in the family Ranidae often go through “boom and bust” cycles from year to year as a result of the environmental instability of their

breeding habitats (Berven 1995). In addition, variance in family size for these frogs may be greater than under random (i.e. Poisson-distributed) expectations due to the loss or survival of whole families during the egg stage of the life cycle (Crow and Morton 1955; Rowe and Beebee 2004). Entire egg masses or portions of egg masses are often lost to desiccation, freezing, predation, or disease (Briggs 1987; McAllister and Leonard 1997). If survival operates at the family level, then the inflation of variance in family size (and reduction in N_b) can be enormous (Crow and Morton 1955). In this study we focus primarily on reduction in N_b incurred during the egg to metamorph stage.

The number of breeding individuals in a given year is often estimated for ranid frog populations by doubling the number of discrete egg masses found in the pond(s) that year (Crouch and Patton, 2000). This estimate is sometimes used to estimate N_b (Merrell 1968; Berven and Grudzien 1990; Watson *et al*. 2000). Estimating N_b this way assumes that each female lays only one egg mass per year, each egg mass is fertilized by a single male, each male breeds with only one female per year, and that family size is Poisson distributed. The first three assumptions are likely to hold for 'explosive breeding' species, which engage in a single, brief (e.g.1-3 nights) reproductive bout each year (Wells 1977). The fourth assumption is much more dubious, but how much reduction in N_b results from non-random survival between egg laying and metamorphosis has not been estimated.

Here we used genetic estimates of N_b in a population of the Oregon spotted frog (*Rana pretiosa*) to estimate the reduction in N_b owing to reproductive strategy and

to non-random survival among families. We analyzed molecular genetic data from large samples of adults, eggs, and post-metamorphic juveniles collected during a single season (Fig. 2.1). We estimated N_b at two stages in the life cycle using variances in microsatellite allele frequencies between: (1) adults and eggs; and (2) adults and metamorphs. This is a single-season version of Waples' (1989) temporal method of N_e estimation, and our approach is similar to that of Scribner et al. (1997).

Given our field observations and the fact that *R. pretiosa* is an explosive breeding species, our *a priori* expectation was that neither extra-pair fertilization nor multiple mating has a strong influence on N_b in this species. Thus, the estimate of the effective number of breeders ($\hat{N}_b$) derived from the allele frequency differences between adults and eggs should be similar to the egg mass count estimate of the actual number of breeders ($\hat{N}_{ab}$ = 2 x number of egg masses). On the other hand, mortality of all or parts of some egg masses is well documented in our and other populations of *R. pretiosa* (Bowerman, personal observation; Licht 1971). Non-random survival among individuals due to egg mass mortality (i.e. family-correlated survival) would reduce $\hat{N}_b$ as measured by allele frequency differences between adults and metamorphs. Therefore, our prediction was that the adult-metamorph $\hat{N}_b$ would be much less than the adult-egg $\hat{N}_b$.

We also estimated $\hat{N}_b$ from eggs and from metamorphs by the linkage disequilibrium (LD) method (Hill 1981). These estimates should be independent of the temporal method estimates (Waples 1991). Again, we predicted that the N_b estimate

from the egg sample would be close to twice the number of egg masses ($\hat{N}_{ab}$), and that the estimate from the metamorph sample would be substantially less than the estimate from the egg sample.

Finally, we estimated N_e (as opposed to N_b) in the adult population via the LD method and compared it to an estimate of N obtained by intensive mark-recapture sampling. These data provide an additional point estimate of N_e/N for ranid frogs.

Materials and Methods

Study Organism

Oregon spotted frogs (*Rana pretiosa*) live in lakes and ponds in the Pacific Northwest, from southern Oregon in the United States to southern British Columbia in Canada (Hayes 1997; Nussbaum *et al*. 1983). *R. pretiosa* overwinter in permanent ponds or springs and breeding occurs soon after ice melt in the spring (Licht, 1969; Leonard *et al*. 1997). During the 2 to 4 week breeding season only mature adults are active at the surface, and the sex ratio is male biased (Watson *et al*. 2000; personal observations). Breeding is explosive, with most of the egg masses being deposited on one or a few nights (Licht 1969; McAllister and Leonard 1997). Females lay their eggs in communal piles in shallow water, and there may be several of these communal sites per pond. Boundaries between egg masses in a pile are very discrete for a week after laying, which makes counting and sampling individual masses straightforward.

The Oregon spotted frog has been extirpated from 70-90% of its original range (Hayes 1997). Fewer than 35 populations remain, and these are mostly small

(N<1000), isolated, and restricted to higher elevations (Hayes 1997; Cushman and Pearl, 2007). *R. pretiosa* is a candidate for federal listing as endangered by the U.S. Fish and Wildlife Service (2005), is considered "sensitive-critical" by the Oregon Department of Fish and Wildlife (Oregon Natural Heritage Information Center 2004), and "endangered" by the state of Washington. It is an endangered species in Canada (Seburn and Seburn 2000). Thus, data on what controls N_e or N_b in this species could be useful for management of the remaining populations.

Sample Collections

Sampling took place in a pond located near Sunriver, Oregon (43.85018° N, 121.44768° W). Adult and post-metamorphic juvenile frogs and were captured using underwater funnel traps (Gee's minnow traps) and dip nets. Adult frogs were individually marked with PIT tags. Metamorphs were not individually marked. Capture-recapture data from marked frogs was collected on 77 occasions from March 6th through December 9th 2006. For genetic sampling, a single toe clip was collected from each adult frog (n = 208) and from each metamorph sampled from the 2006 cohort (n = 401). Toe-clips were stored in Drierite desiccant (W. A. Hammond Drierite Co., Xenia, OH). During the breeding season (late March through early April), the pond was carefully monitored for the presence of egg masses. 45 egg masses were deposited on April 6th and were sampled within 48 hrs. We observed no egg mass mortality prior to taking our egg samples. Approximately 10 eggs were sampled from each mass (n = 452). The eggs were allowed to develop for several days

in the laboratory and then preserved in 70% ethanol. To our knowledge, no additional egg masses were deposited in 2006 and thus our sample of eggs included all families for that year. We excluded from our datasets any individual with missing data for one or more microsatellite loci. This exclusion resulted in an adult sample of 176, a total egg sample of 415, and a metamorph sample of 308.

The methods of N_b estimation used in this study assume samples are drawn at random (Hill 1981; Waples 1989). By collecting roughly 10 eggs from each egg mass, we may have forced allele frequencies estimated from the egg sample to be more similar to the adult frequencies than if the same number of eggs had been sampled randomly from the entire pool of eggs produced in the pond (Waples, personal communication). This imposed uniformity could result in an upward bias of the N_b estimates obtained using the egg sample. To avoid this potential bias, we generated a corrected sample of eggs by drawing a random number of individuals from each egg mass (using a Poisson distribution with $\lambda = 4$; random numbers from this distribution ranged from 0-10). We generated five of these corrected samples with replacement (n ranged from 156-183), estimated N_b separately for each (see methods below), and then calculated the harmonic mean of $\hat{N}_b$ across the five samples. We report the mean, bias-corrected $\hat{N}_b$ values, though we found that these were very similar to the values obtained using the entire sample of ~10 eggs per mass.

Microsatellite genotyping and scoring

Total genomic DNA was extracted from each sample using QIAGEN DNeasy kits (QIAGEN Inc.). Each individual was genotyped at 7 microsatellite loci (Table 2.1). PCR amplifications were run in 20 μl volumes with the following components: 100-200 ng genomic DNA, 25 mM KCl, 1 mM Tris-HCl pH 9, 0.1% Triton X-100, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.5 μM both forward (fluorescently-labeled) and reverse primers, 0.01 units/μl of Taq, and water to a final volume of 20 μl. PCR amplifications were carried out in an MJ Research PTC-200 thermal cycler under the following conditions: 94° C for 3 min, followed by 30 cycles of 94° C for 30 s, locus-specific annealing temperature (Table 2.1) for 30 s, 72° C for 30 s, and a final extension of 72° C for 7 min. Microsatellite PCR products were run on an ABI 3730 automated sequencer, and allele sizes were scored using the program GENOTYPER v. 3.7 (Applied Biosystems). The program GENEPOP (Raymond and Rousset 1995) was used to estimate allele frequencies and to test loci for deviations from Hardy-Weinberg equilibrium. We tested all pairs of loci for linkage disequilibrium using the program FSTAT 2.9.3 (Goudet 2002).

Estimates of census population size

We had extensive mark-recapture data for 2006 season, which allowed us to obtain estimates of the adult population size ($\hat{N}$) using Begon's weighted mean method (Begon 1979) and the program CAPTURE (White *et al*. 1978). Begon's weighted mean is a modification of the simple Lincoln-Peterson estimate that utilizes capture data from >2 trapping occasions. CAPTURE uses maximum likelihood and a

model selection procedure to identify the model that best fits the mark-recapture data from among eight possible models. The eight models differ in what variables they include: effects of time on capture, behavioral effects (e.g. "trap-happy" or "trap-shy" behaviors) on capture, and individual variation in capture probability (White *et al*. 1978). We averaged the estimated population size from the best-fitting model identified by CAPTURE with the estimate obtained using Begon's weighted mean.

Estimates of N_b *and* N_e

There are several methods of estimating N_e indirectly using genetic data, and the time frame to which an estimate applies depends on the method used as well as the sampling design (Waples 2005). In any case, N_e applies to one or more *generations*, whereas N_b is the effective number of breeding adults in a *single reproductive season* that produce a single cohort of offspring. N_e is difficult to derive from N_b for organisms with overlapping generations because this requires extensive demographic information about the population (Jorde and Ryman 1995; Waples 2005). However, low estimates of N_b are generally expected to reflect low N_e (Waples 2005). N_b can be estimated by the same methods used to estimate the overall effective size of a population.

Although the 13 microsatellite loci we developed for *R. pretiosa* (Blouin, unpublished data) were polymorphic when surveyed across the species' range (Blouin 2002), only 7 proved to be polymorphic in the Crosswater population and none had more than 3 alleles at a locus. This low level of genetic diversity precluded the use of

kinship and pedigree methods to accurately match offspring with their parents or siblings, preventing direct, pedigree-based estimation of N_b (e.g. Araki *et al*., 2007; Blouin, 2003). Consequently, we estimated N_b only through indirect genetic methods.

We compared the estimate of the actual number of breeders obtained from an egg mass count ($\hat{N}_{ab}$) to several $\hat{N}_b$ values obtained from genetic data. The first $\hat{N}_b$ is from a version of the temporal method that uses the differences in allele frequencies between a sample from the adult population and a sample from their offspring (Scribner *et al*. 1997). We calculated two $\hat{N}_b$ values: one based on allele frequency differences between the adult and egg samples, and another based on allele frequency differences between the adult and metamorph samples. As noted above, each adult-egg estimate of N_b that we report represents the harmonic mean of five random samples generated from the total egg dataset. See Table 2.2 for notation used.

The first temporal method we used was Waples' (1989) moment-based approach (TM). The standardized variance of allele frequency change for each locus was calculated using equation (9) from Waples (1989):

$$\hat{F}_c = \frac{1}{K-1}\sum_{i=1}^{K}\frac{(x_i - y_i)^2}{(x_i + y_i)/2} \qquad (1)$$

where K is the total number alleles at the locus, x_i is the frequency of allele i in the first sample, and y_i is the frequency in the second sample. The mean $\hat{F}_c$ across all 7 loci was calculated as:

$$\text{mean } \hat{F}_c = \sum (K_j - 1) F_{c_j} / \sum (K_j - 1) \qquad (2)$$

where K_j is the number of alleles at locus *j* and F_{c_j} is the estimate of F_c for locus *j*. Confidence intervals for mean $\hat{F}_c$ were calculated using equation (16) from Waples (1989). Because our first sample was collected non-destructively from adults, Waples' (1989) Plan I was the appropriate sampling design. The estimated effective number of breeders was therefore calculated using equation (12) from Waples (1989):

$$\hat{N}_b = \frac{t}{2\left[\text{mean } \hat{F}_c - \frac{1}{2S_o} - \frac{1}{2S_t} + \frac{1}{\hat{N}}\right]} \qquad (3)$$

where S_o and S_t are sample sizes for the first and second samples, respectively, *t* is number of generations between the two samples (1 in this case), and $\hat{N}$ is the census estimate of the total size of the population from which the S_o sample was drawn (see above for how we obtained $\hat{N}$). We designated the adult-egg and adult-metamorph estimates of N_b from this method as $\hat{N}_b^{TM-Egg}$ and $\hat{N}_b^{TM-Meta}$, respectively.

The second temporal approach we used to estimate N_b was the likelihood-based estimator (TL) of Berthier *et al*. (2002), implemented in the program TM3. The TL method involves the calculation of likelihoods from coalescent-based gene genealogies and Markov chain Monte Carlo sampling to generate a posterior

probability distribution of N_e, or in our case, N_b. We designated the adult-egg and adult-metamorph estimates of N_b from this method as $\hat{N}_b^{TL-Egg}$ and $\hat{N}_b^{TL-Meta}$, respectively. A maximum possible N_b value is specified as a Bayesian prior in TM3. Although we did not expect maximum $\hat{N}_b$ to be greater than about 90 frogs (based on the egg mass count), we ran several independent TM3 runs using priors of 200, 300, 400, and 1000 for maximum N_b. We set our lowest prior conservatively at 200 to account for the possibility that extra-pair fertilization (i.e. multiple fathers per egg mass) could result in N_b greater than $\hat{N}_{ab}$. Performing analyses with different priors allowed us to evaluate the sensitivity of $\hat{N}_b^{TL-Egg}$ and $\hat{N}_b^{TL-Meta}$ to choice of prior. All TM3 analyses were run with 50,000 iterations.

In addition to the two temporal methods, we used the linkage disequilibrium (LD) method to estimate N_b from single samples of eggs and of metamorphs. We designated the N_b estimates from the LD method as $\hat{N}_b^{LD-Egg}$ for the egg samples and as $\hat{N}_b^{LD-Meta}$ for the metamorph sample. Calculations were performed using the program LDNe (Waples and Do 2007). LDNe incorporates a correction for the bias that is introduced when sample size is less than the true effective size and reports confidence intervals obtained via a new jackknife method (Waples 2006). The mating model for this system is equivalent to monogamy and was selected in the LDNe analyses. We report jackknife confidence intervals for $\hat{N}_b$, with the lowest allele frequency set at 0.05. By excluding alleles with frequencies less than 0.05, we achieve the most accurate N_b estimate, with an expected tradeoff in precision (Waples and Do

2007). However, even when we ran our analyses with the lowest allele frequency set at 0.01, confidence intervals were very similar to those obtained when the lowest frequency was set at 0.05.

Lastly, we used the LD method to estimate N_e (as opposed to N_b) in the adult sample (n = 176), under a random mating model in LDNe. We acknowledge that there is some uncertainly about how to interpret LD estimates from mixed-cohort samples from species that have overlapping generations (Waples 1991). However, the LD method has become standard for estimating N_e from such samples (e.g. Aspi et al 2008; Durrant et al 2008), so our data should still be useful for comparative purposes.

Results

Genetic diversity

Expected heterozygosity (H_e) for the 7 microsatellite loci in this *R. pretiosa* population was 0.40 as calculated from the adult sample. The maximum number of alleles per locus was 3. Only one locus in one sample (RP3 in the metamorph sample) was barely out of Hardy-Weinberg equilibrium (P = 0.0071; Bonferroni-corrected nominal value of 0.00714). We found one locus pair (RP3 x RP385) with barely significant linkage disequilibrium in the adult sample (P = 0.00238; Bonferroni-corrected nominal value of 0.002381). Six pairs of loci in the metamorph sample and 3-4 pairs in each of the 5 random egg samples exhibited significant linkage disequilibrium (data not shown).

Estimates of census population size

The two methods of population size estimation yielded very similar results. Begon's weighted mean method gave $\hat{N}$ of 444 (95% C.I. 343-545). CAPTURE identified the M_t model as the most appropriate for our mark-recapture data. Under this model, each individual has the same probability of capture on a given trapping occasion, but these probabilities are variable across trapping occasions (White *et al*. 1978). $\hat{N}$ from the CAPTURE analysis was 412 (95% C.I. 343-513). The average of the two $\hat{N}$ values is 428 (95% C.I. 343-529).

Estimates of N_b and N_e

Point estimates of N_b from the TL analysis using the program TM3 were insensitive to the value of Bayesian prior for maximum N_b (Table 2.3). As one might expect, the upper confidence limit did increase with increasing prior. However, even if extra-pair fertilization was rampant in this population, such that the number of breeding males was more than twice the number of breeding females, maximum N_b should not exceed 200. Thus, using 200 as the upper prior for our reported values (Table 2.4) probably produced overly liberal upper confidence intervals, even if we are confident in the point estimates.

Estimates of N_b are presented in Table 2.4. Doubling the number of egg masses found in the 2006 breeding season resulted in an estimate of 90 breeding adults ($\hat{N}_{ab}$ = 90). The temporal methods (TM and TL) yielded similar point estimates of N_b for the

adult-egg ($\hat{N}_b^{TM-Egg} = 65.0$, $\hat{N}_b^{TL-Egg} = 87.3$) and adult-metamorph comparisons ($\hat{N}_b^{TM-Meta} = 82.5$, $\hat{N}_b^{TL-Meta} = 117.2$). Estimates of N_b from the LD method were also very similar between the two life stages ($\hat{N}_b^{LD-Egg} = 68.5$, $\hat{N}_b^{LD-Meta} = 56.2$). Thus, we see (1) point estimates from the egg stage (65.0, 87.3 and 68.5) that are fairly close to the simple estimate of 90 breeding adults, and (2) no indication of a massive drop in N_b in going from the egg to metamorph stage.

The LD estimate of N_e in the adult sample was 36.7 (95% C.I. 19-71.9). Thus, the best point estimate of $N_{e/} N$ for this population = 36.7/428 = 0.086.

Discussion

Esitmates of N_b/N across the various methods ranged from 0.13 to 0.27 (Table 2.4). These values are similar to those found for *R. temporaria* populations in Finland (0.06-0.17; Schmeller and Merila 2006) and Britain (0.333-0.365; Brede and Beebee 2006), but considerably higher than those of toad (*Bufo bufo*) populations in Britain (0.007-0.012 Scribner *et al*. 1997; 0.034-0.040 Brede and Beebee, 2006). The N_e/N ratio estimated for the adult population was 0.086, which again is in the general range of DNA-based estimates for other ranid frogs (Hoffman et al 2004; Schmeller and Merila 2006). Thus, our data are consistent with previous suggestions that N_e/N ratios in ranid frogs are in the typical range for vertebrates (e.g. ~0.1 to 0.4), while those for bufonids are much lower (Hoffman *et al*. 2004; Brede and Beebee 2006).

Our main objective was to test a hypothesis about what features of the life cycle of *R. pretiosa* cause N_e to be reduced relative to N. By obtaining separate N_b estimates using egg and metamorph samples we could determine if these N_b estimates differed from each other and from the simple estimate from counting egg masses ($\hat{N}_{ab}$). To our knowledge, this study is the first to take such an approach. We found that: (1) Estimates for N_b at the egg stage using both temporal methods and the LD method did not differ dramatically from $\hat{N}_{ab}$ = 90; and (2) estimates for N_b were similar for eggs and metamorphs (Table 2.4). The first result is consistent with what we would expect to find if each female produced a single egg mass, each egg mass was fertilized by a single male, and each male bred with only one female. The second result suggests little non-random (family-based) mortality occurred between egg laying and metamorphosis.

The first result was expected because, like many ranid frogs, *R. pretiosa* females are thought to lay one egg mass per season (Olson and Leonard 1997) and explosive breeding reduces the opportunity for males to mate with multiple females (Wells, 1977). Indeed, in this year all breeding occurred on a single night. On the other hand, sex ratios in breeding populations are male-biased, which could promote extra-pair fertilization, so that some egg masses are fertilized by more than one male. This multiple paternity could occur either passively by free-swimming spermatozoa in communal breeding areas (Laurila and Seppa 1999) or actively by lone 'pirate' (or 'sneaker') males that fertilize some of the eggs of breeding pairs (Vieites et al 2004). If extra-pair fertilization was frequent, the resulting decrease in the variance of male

reproductive success would increase N_b estimates at the egg stage relative to $\hat{N}_{ab}$ (Sugg and Chesser 1994). In our observations of hundreds of breeding pairs of spotted frogs over multiple years in this population, we have witnessed few instances of behavior that would suggest the occurrence of 'clutch piracy.' Thus, we interpret our results as consistent with predictions of a basically monogamous mating system in which each female lays a single clutch per year. One practical consequence of these results is that they support the use of egg mass counts as a cost-effective method of population monitoring, in that they probably do give a reasonable estimate of the number of adults that bred in a given year. Whether egg mass counts can consistently provide reliable estimates of N_b depends on how typical are the results that variance in family size apparently increased little after egg laying.

If our results for this population in 2006 are typical for *R. pretiosa*, then the reduction of *Ne* relative to *N* in this species is not owing to the inflation of variance in family size that occurs between the egg and metamorph stages. Thus, we might consider other factors such as year-to-year fluctuations in population size. Of course, our results are from one year in a single population and may not be typical. Water levels in the pond were very high in 2006, which may have contributed to unusually high survival of entire egg masses. Such an environmental effect was also noted by Schmeller and Merila (2007), who suggested that high egg-to-metamorph mortality during a short growing season may have been responsible for low N_b/N ratios in two populations of *Rana temporaria.* Together, these observations suggest the interesting hypothesis that the N_b/N ratio varies from year to year (or from population to

population) depending on habitat quality. Indeed, N_b/N and N_e/N might even be predictable from environmental measurements.

This study is the first attempt to determine where in the ranid life cycle N_b (and by extension, N_e) is reduced. More studies will be needed before a consensus is reached about the importance of different factors. Here we provide some of the first data on the subject, and suggest the hypothesis that N_b/N might vary substantially in time and space owing to habitat conditions that influence the survival of eggs and larvae. The approach of estimating N_b using genetic data from a single cohort at more than one life stage should prove valuable in future studies on the determinants of effective population size in amphibians and other taxa.

Acknowledgements

We would like to thank Patty Stenberg for help in tissue collection. Robin Waples provided suggestions on the methodology and helpful comments on an early draft of the manuscript. Two anonymous reviewers also provided useful input on the manuscript. We thank the Oregon State University Center for Genome Research and Biocomputing (CGRB) for assistance with microsatellite genotyping.

Literature Cited

Araki H, Waples RS, Ardren WR, Cooper B, Blouin MS (2007) Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms. Mol Ecol 16:953-966

Aspi J, Roininen E, Kiiskilä J, Ruokonen M, Kojola I, Bljudnik L, Danilov P, Heikkinen S, Pulliainen E (2008) Genetic structure of the northwestern Russian wolf populations and gene flow between Russia and Finland. Conservat Genet doi:10.1007/s10592-008-9642-x.

Beebee T (2005) Conservation genetics of amphibians. Heredity 95:423-427

Begon M (1979) Investigating animal abundance: capture-recapture for biologists University Park Press, Baltimore, MD.

Berthier P, Beaumont MA, Cornuet J-M, Luikart G (2002) Likelihood-Based Estimation of the Effective Population Size Using Temporal Changes in Allele Frequencies: A Genealogical Approach. Genetics 160:741-751

Berven KA, Grudzien TA (1990) Dispersal in the wood frog (*Rana sylvatica*): Implications for genetic population structure. Evolution 44, 8:2047-2056

Berven KA (1995) Population regulation in the wood frog, *Rana sylvatica*, from three diverse geographic localities. Aust J Ecol. 20:385–392

Blouin MS (2002) Genetic data for recovery of the Oregon Spotted Frog, *Rana pretiosa*, in the Western United States, Final report for USGS FY 2000 Species at Risk Program project. USGS.

Blouin MS (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends Ecol Evol 18:503-511

Brede EG, Beebee TJ (2006) Large variations in the ratio of effective breeding and census population sizes between two species of pond-breeding anurans. Biol J Linn Soc 89:365-372

Briggs JL (1987) Breeding Biology of the Cascade Frog, *Rana cascadae*, with Comparisons to *R. aurora* and *R. pretiosa*. Copeia 1987:241-245

Crouch WB, Paton PWC (2000) Using egg-mass counts to monitor wood frog populations. Wildlife Wildl Soc Bull 28:895-901

Crow JF, Morton NE (1955) Measurement of gene frequency drift in small populations. Evolution 9:202-214

Cushman KA, Pearl CA (2007) A conservation assessment for the Oregon spotted frog (*Rana pretiosa*). USDA Forest Service and USDI Bureau of Land Management, Oregon and Washington.

Durrant C, Beebee T, Greenaway F, Hill D (2008) Evidence of recent population bottlenecks and inbreeding in British populations of Bechstein's bat, *Myotis bechsteinii*. Conservat Genet doi:10.1007/s10592-008-9639-5. http://dx.doi.org/10.1007/s10592-008-9639-5

Easteal S (1985) The ecological genetics of introduced populations of the giant toad *Bufo marinus*. II. Effective population size. Genetics 110:107-122

England P, Cornuet J, Berthier P, Tallmon D, Luikart G (2006) Estimating effective population size from linkage disequilibrium: severe bias in small samples. Conservat Genet 7, 303-308.

Frankham R (1995) Effective population size/adult population size ratios in wildlife: A review. Genet Res 66:95-107

Funk WC, Blouin MS, Corn PS, *et al.* (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. Mol Ecol 14, 483-496.

Goudet J (2002) FSTAT version 2.9. 3.2, updated from Goudet (1995) FSTAT: A computer program to calculate F-statistics. J Hered 86:485-486.

Hayes MP (1997) Status of the Oregon spotted frog (*Rana pretiosa*) in the Deschutes Basin and selected other systems in Oregon and northeastern California with a rangewide synopsis of the species' status. Nature Conservancy under contract to the U.S. Fish and Wildlife Service with assistance from the Oregon Department of Fish and Wildlife, final report, Portland, Oregon.

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. Genet Res 38:209-216.

Hoffman EA, Schueler FW, Blouin MS (2004) Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern leopard frog. Evolution 58:2536-2545.

Jorde PE, Ryman N (1995) Temporal Allele Frequency Change and Estimation of Effective Size in Populations with Overlapping Generations. Genetics 139:1077-1090.

Laurila A, Seppa P (1998) Multiple paternity in the common frog (*Rana temporaria*): genetic evidence from tadpole kin groups. Biol J Linn Soc 63:221-232.

Leonard W, Hallock L, McAllister KR (1997) Natural history notes: *Rana pretiosa* (Oregon spotted frog). Behavior and reproduction. Herpetological Review 28:86.

Licht L (1969) Comparative breeding behavior of the red-legged frog (*Rana aurora aurora*) and the western spotted frog (*Rana pretiosa pretiosa*) in southwestern British Columbia. Can J Zool 47:1287-1299

Licht LE (1971) Breeding Habits and Embryonic Thermal Requirements of the Frogs, *Rana Aurora Aurora* and *Rana Pretiosa Pretiosa*, in the Pacific Northwest. Ecology 52:116-124

Merrell DJ (1968) A Comparison of the estimated size and the "effective size" of breeding populations of the leopard frog, *Rana pipiens*. Evolution 22:274-283

McAllister KR, Leonard WP (1997) Washington State status report for the Oregon Spotted Frog, p. 38. Washington Department of Fish and Wildlife, Olympia.

Monsen KJ, Blouin MS (2003) Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. Mol Ecol 12:3275-3286

Nussbaum RA, Brodie Jr. ED, Storm RM (1983) *Amphibians and Reptiles of the Pacific Northwest* University of Idaho Press, Moscow, Idaho.

Olson DH, Leonard WP (1997) Amphibian inventory and monitoring: a standardized approach for the PNW. In: *Sampling Amphibians in Lentic Habitats* (eds. Olson DH, Leonard WP, Bury RB), pp. 1-21. Society for Northwestern Vertebrate Biology, Olympia, WA.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. J Hered 86:248-249

Schmeller DS, Merila J (2007) Demographic and Genetic Estimates of Effective Population and Breeding Size in the Amphibian *Rana temporaria*. Conservat Biol 21:142-151

Scribner KT, Arntzen JW, Burke T (1997) Effective number of breeding adults in *Bufo bufo* estimated from age-specific variation at minisatellite loci. Mol Ecol 6:701-712

Seburn D, Seburn C (2000) Conservation priorities for the amphibians and reptiles of Canada. Prepared for World Wildlife Fund Canada and the Canadian Amphibian and Reptile Conservation Network.

Stuart SN, Chanson JS, Cox NA, *et al.* (2004) Status and Trends of Amphibian Declines and Extinctions Worldwide. Science 306:1783 - 1786

Sugg DW, Chesser RK (1994) Effective population size with multiple paternity. Genetics 137:1147–1155

US Fish and Wildlife Service (2005) Species assessment and listing priority assessment form. Western Washington Fish and Wildlife Office, Lacey, Washington. Available at http://ecos.fws.gov/docs/candforms_pdf/r1/D02A_V01.pdf

Vieites D, Nieto-Román S, Barluenga M, *et al.* (2003) Post-mating clutch piracy in an amphibian. Nature Rev. Genet 4:598-612

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics 121:379-391

Waples RS (1991) Genetic methods for estimating the effective size of cetacean populations. Report of the International Whaling Commission (special issue) 13:279–300

Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? Mol Ecol 14:3335-3352

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservat Genet 7:167-184

Waples RS, Do C (2007) LDNE: a program for estimating effective population size from data on linkage disequilibrium. Mol Ecol Resources 8:753-756

Watson JW, McAllister KR, Pierce DJ, Alvarado A (2000) Ecology of a remnant population of Oregon spotted frogs (*Rana pretiosa*) in Thurston County, Washington. Washington Department of Fish and Game, Olympia, WA.

Wells K (1977) The social behaviour of anuran amphibians. Anim Behav 25:666-693

White GC, Burnham KP, Otis DL, Anderson DR (1978) *Users manual for program CAPTURE* Utah State University Press, Logan, UT.

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97-159

Table 2.1. Microsatellite primer information.

Locus	F primer R primer	Annealing temp (°C)	# of alleles
RP3[†]	5'gaaagcaaaactgggaaagtacata3' 5'cctgagagccatccaataagtgcca3'	50	3
RP22	5'accccaccagcagaatacaatga3' 5'agaccagagccagagcaacc3'	50	3
RP23[†]	5'acatagatacaatagatagatagac3' 5'cacaggaatgtaaaatctggctttc3'	45	3
RP193*	5'ccattttctctctgatgtgtgt3' 5'tgaagcagatcactggcaaagc3'	50	2
RP385	5'attgaaacttgcggctctct3' 5'ggcatgtgtccacaatgtaa3'	50	2
RP415	5'aagtttcattaaagcagatt3' 5'ggtatatcttagggttacct3'	45	2
SFC134*[†]	5'tgggaaaagactctgtggt3' 5'aggaaatgtgtggaagcat3'	55	3

* Locus used in Monsen and Blouin (2003)
[†] Locus used in Funk et al. (2005)

Table 2.2. Notation used.

N	Actual number of individuals in the population; population census size
N_{ab}	actual number of breeding adults
N_e	Effective population size
N_b	Effective number of breeding adults in one reproductive season
$\hat{N}$, $\hat{N}_{ab}$, $\hat{N}_e$, $\hat{N}_b$	Estimates of N, N_{ab}, N_e, and N_b
TM	Waples' (1989) temporal moment method of estimating N_e or N_b
TL	Temporal likelihood method of N_e (N_b) estimation from Berthier *et al*. (2002)
LD	Linkage disequilibrium method of N_e (N_b) estimation
$\hat{N}_b^{TM-Egg}$	Temporal moment method estimate of N_b, using the adult and egg samples
$\hat{N}_b^{TM-Meta}$	Temporal moment method estimate of N_b, using the adult and metamorph samples
$\hat{N}_b^{TL-Egg}$	Temporal likelihood method estimate of N_b, using the adult and egg samples
$\hat{N}_b^{TL-Meta}$	Temporal likelihood method estimate of N_b, using the adult and metamorph samples
$\hat{N}_b^{LD-Egg}$	LD method estimate of N_b from LDNE program, using the egg sample
$\hat{N}_b^{LD-Meta}$	LD method estimate of N_b from LDNE program, using the metamorph sample

Table 2.3. Harmonic means of $\hat{N}_b$ values for four choices of Bayesian prior for maximum N_b. Means were calculated from the 5 bias-corrected egg samples for each prior.

Prior	Harmonic Mean	lower C.L.	upper C.L.
200	86.35	20	200
300	83.37	17	280
400	91.04	19	352
1000	88.61	15	519

Table 2.4. Estimates of effective number of breeders (N_b) and N_b/N in the CW population. Estimates are given for two temporal methods: Waples' (1989) temporal moment (TM) and the temporal likelihood method (TL) of Berthier *et al*. (2002). Estimates from the linkage disequilibrium method were obtained using LDNE (Waples and Do 2007). Estimates from these various methods are listed along with their 95% confidence intervals (Bayesian credible intervals for the TL estimates) and N_b/N ratios.

Method			Estimate	95% C.I.	N_b/N
Temporal methods					
Adult-egg:					
	TM	$\hat{N}_b^{TM-Egg}$	65.0	18-195	0.15
	TL	$\hat{N}_b^{TL-Egg}$	86.4	20-200	0.20
Adult-metamorph:					
	TM	$\hat{N}_b^{TM-Meta}$	82.5	23-252	0.19
	TL	$\hat{N}_b^{TL-Meta}$	117.2	27-200	0.27
Linkage Disequilibrium methods					
Egg:					
	LD	$\hat{N}_b^{LD-Egg}$	68.5	30-108	0.16
Metamorph:					
	LD	$\hat{N}_b^{LD-Meta}$	56.2	26-108	0.13

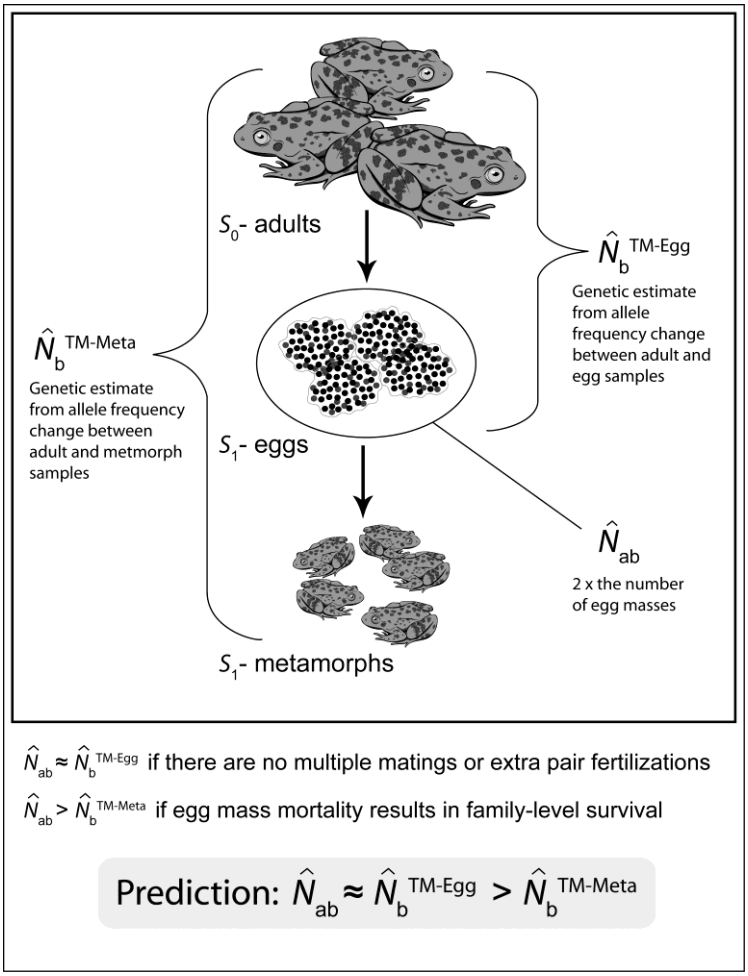

Figure 2.1. Sampling scheme for estimating the effective number of breeders (N_b) in a population of *R. pretiosa*. Three samples were collected in 2006: adults, eggs, and metamorphs. For each sample, allele frequencies were calculated for 7 microsatellite loci. Two estimates of N_b were derived from this genetic data using the temporal method. The first was based on allele-frequency differences between the adult and egg samples ($\hat{N}_b^{TM-Egg}$), while the second was based on differences between the adult and metamorph samples ($\hat{N}_b^{TM-Meta}$). An estimate of the actual number of breeding adults was calculated as twice the number of egg masses counted in the pond in 2006 ($\hat{N}_{ab}$). Our predictions as described in the text are represented by the relationships among $\hat{N}_b^{TM-Egg}$, $\hat{N}_b^{TM-Meta}$, and $\hat{N}_{ab}$. Note that the 'TM' superscript used here refers to Waples' (1989) temporal moment method, but this sampling scheme and data were also used for the temporal likelihood (Berthier *et al*. 2002) analysis. Estimates of N_b were obtained separately for the egg and metamorph samples using the LD method (Hill 1981; Waples and Do 2007).

CHAPTER 3

Effective size of ranid frog populations: comparisons within and among four species

Ivan C. Phillipsen, W. Chris Funk, Eric A. Hoffman, Kirsten J. Monsen, and Michael S. Blouin

Abstract

We used microsatellite loci to estimate effective size (N_e) in each of 90 populations of four species of ranid frogs (20 to 26 populations per species, mean *n* per population = 29). Our objectives were to (1) determine typical values of N_e for populations of each species, (2) determine the strength of the correlation between estimates of genetic diversity and N_e among populations in each species, (3) compare N_e estimates among the species, and (4) test for correlations between each of several geographic variables and N_e estimates within species. We used single-sample linkage disequilibrium (LD) and approximate Bayesian computation (ABC) methods to estimate contemporary N_e for each population. We compared these estimates with temporal method estimates for 7 populations. Three of the species in our comparison—*Rana pretiosa, R. luteiventris*, and *R. cascadae*— have consistently small effective population sizes (<50) and low genetic diversities. In contrast, N_e in *Lithobates (Rana) pipiens* spans a much wider range, with many values of N_e in the hundreds. There was a strong correlation between genetic diversity and N_e. Thus, genetic diversity appears to be a good proxy for recent effective size in ranid frogs. Estimates from the LD and ABC methods showed significant, positive correlations within populations of only *R. luteiventris* and *L. pipiens*. We also found significant correlations between N_e estimates and latitude, longitude, or altitude in these two species. We discuss some hypotheses to explain the differences in N_e among species and the correlations between N_e and various environmental variables within species.

Introduction

A key parameter in the theory and application of population genetics is the effective population size (N_e): the number of breeding individuals in a conceptual, ideal population that would lose genetic diversity at the same rate as the real population being studied (Wright 1931; Charlesworth 2009). How a population responds to evolutionary forces depends on N_e, rather than the actual number of individuals in the population (N, the census population size). Although direct estimates of N_e can be calculated from demographic data, such data are often prohibitively difficult to obtain (Wang 2005). Given this limitation, and the importance of N_e in population and conservation genetics, it is not surprising that considerable effort has been put into developing methods of using molecular genetic data to obtain indirect estimates of N_e.

With advances in these methods and with the increasing accessibility of multilocus genotype data it has recently become practicable to estimate the effective sizes of many populations of the same species (Luikart *et al.* 2010; Waples and Do, 2009). Such an effort is highly worthwhile for several reasons. First, by gathering estimates from multiple populations, investigators might identify a reasonably narrow range of typical N_e values for a species, an "educated guess" for the value of this parameter in any given population. This information could be used to approximate N_e in evolutionary modeling of the species. For example, an expected N_e could be used as input for simulations of evolutionary processes or as a Bayesian prior in analyses used to infer other population genetic parameters, such as migration rates or selection

coefficients. Expected N_e values for a species would also be useful in conservation and management, in situations where an estimate of N_e for a population is desired but genetic and demographic data are unavailable. In such cases it would be helpful to know if populations of the focal species typically have, for example, N_e estimates less than 100. Furthermore, when estimates of both N_e and N (census size) can be obtained for a large number of populations, it may be possible to identify the typical range of N_e/N for the species. If so, estimates of N for the species might then be used as proxies for N_e, when the former are easier to obtain that the latter.

Second, estimates of N_e from multiple populations can be used to investigate the correlation between N_e and genetic diversity (e.g. expected heterozygosity, H_e) within populations, which in turn is important because H_e is easier to estimate than N_e, and estimates of H_e are abundant in older literature. Thus, it would be useful to know if variation among populations in H_e could be used as a proxy for variation in N_e. Although a positive correlation between N_e and H_e is expected at equilibrium (Soulé 1976; Frankham 1996), this relationship may not always hold and can vary among taxa. For example, H_e might not be a reliable proxy for N_e in species whose populations vary substantially in their degree of connectivity (via gene flow) or that tend to fluctuate in size. The extent to which estimates of H_e serve as useful proxies for estimates of N_e remains to be tested for most taxa.

Third, a comparative analysis of N_e estimates from multiple populations across more than one species can help generate hypotheses about what particular biological factors influence N_e within the species. Consistent differences in N_e among species

might correspond to differences in habitat, dispersal capabilities, or breeding behaviors. Hypotheses generated in a comparative analysis could then be tested in subsequent studies. Similarly, the factors that influence N_e within species can also be investigated by evaluating correlations between environmental variables and N_e for multiple populations of a species.

Despite the fact that such multi-population, empirical investigations of N_e can lead to valuable insights, surprisingly few examples of this approach exist (Fraser *et al*. 2007; Beebee 2009). In this study, we employed an unusually large dataset of 90 populations to conduct a comparative analysis of N_e within and among four species of North American frogs in the family Ranidae: the Oregon spotted frog (*Rana pretiosa*), the Columbia spotted frog (*Rana luteiventris*), the Cascades frog (*Rana cascadae*), and the northern leopard frog (*Lithobates* [*Rana*] *pipiens*). Although the focus of our discussion is on these species, the methods we used should be applicable in many other study systems. Thus, this study demonstrates the general value of gathering and analyzing N_e estimates from multiple populations and species, in addition to furthering our knowledge of N_e in ranid frogs.

We described the population structures of the four frog species in previous studies that did not explicitly assess N_e (Blouin *et al., in press*); Funk *et al.* 2005; Funk *et al.* 2008; Monsen and Blouin 2003; Monsen and Blouin 2004; Hoffman and Blouin 2004a and 2004b). Research on other species suggests that amphibian populations tend to have small N_e, on the order of a few 10s to no more than a few thousand (e.g. Easteal 1985; Funk *et al.* 1999; Jehle *et al.* 2001; Brede and Beebee 2006; Schmeller

and Merila 2007; Beebee 2009). Still, much remains unknown about N_e in amphibians, as trends in N_e estimates across multiple populations have not been analyzed for the vast majority of species.

We used the single-sample linkage disequilibrium (LD; Hill 1981) method and a newer, approximate Bayesian computation method (ABC; Tallmon *et al.* 2008) to estimate contemporary N_e for each population. Estimates from these methods apply to N_e over a time scale spanning a few generations at most (Waples 2005). The ABC method incorporates more information than the LD method and so should be more accurate. But the LD method is easier to apply and has been in use for many years, so it is worth knowing to what extent the two methods give similar results. Therefore, we assessed the correlation between the LD and ABC estimates within each species. We also compared these single-sample estimates with estimates from the temporal method (Waples 1989) for two of the *R. pretiosa* populations and five of the *L. pipiens* populations.

Our main objectives in this study were to: (1) determine the typical N_e values for each species, (2) determine the strength of the correlation between genetic diversity and N_e estimates, (3) compare N_e estimates among the species and offer hypotheses to explain the differences, and (4) test for correlations between each of several geographic variables and N_e estimates within each species. Patterns in N_e estimates at these levels of comparison may reflect important population genetic processes in these frogs, knowledge of which would be valuable to the study of

amphibian evolutionary biology and to the conservation management of these organisms.

Materials and Methods

Study species

Frogs in the family *Ranidae* are distributed widely at the global scale, with 28 species occurring in the United States and Canada (Collins and Taggart 2009). *Rana pretiosa*, *R. luteiventris*, and *R. cascadae* are distributed across the northwesten region of North America. *R. pretiosa* and *R. cascadae* have relatively small ranges that overlap in Oregon and Washington; *R. luteiventris* has a larger range that extends from Utah to Alaska. Most populations of these 'northwestern' species are found in lakes and ponds in mountain environments (Jones *et al.* 2005). *R. pretiosa* and the Great Basin populations of *R. luteiventris* are candidates for federal listing as endangered in the United States (U.S. Fish and Wildlife Service 2005; U.S. Fish and Wildlife Service 2004), and *R. cascadae* is a Species of Concern at the federal level. Each of the three species has some level of protection at the state or provincial level. *R. pretiosa* is listed as endangered in Canada (Seburn and Seburn 2000).

The natural range of *L. pipiens* is one of the largest for a North American amphibian, spanning much of the continent (Stebbins 2003). The primary habitat of *L. pipiens* is valley wetlands and forests (Rorabaugh 2005). Western populations of this species are currently under review for federal listing as endangered in the United States (Federal Register 2009).

Sampling and Molecular Methods

We collected population samples and obtained genetic data from the four frog species as described in previous studies (Fig. 3.1, Table A.1): 21 *R. pretiosa* populations were sampled from across the species' range in Oregon and Washington (Blouin 2000); 26 *R. luteiventris* populations were sampled in the mountains of Montana and Idaho (Funk *et al.* 2005); 20 *R. cascadae* populations were sampled at a similar geographic scale, across the Cascades Mountains of Oregon and Washington (Monsen and Blouin 2003); and 23 *L. pipiens* populations were sampled in the Canadian provinces of British Columbia and Ontario and in the states of Idaho, Nebraska, and Minnesota (Hoffman and Blouin 2004b; Hoffman *et al.* 2006). Two *R. pretiosa* populations and 5 *L. pipiens* populations are each represented by two temporally-spaced samples (Table A.1). Populations of these species are likely isolated by low gene flow, as indicated by their strong genetic differentiation at small spatial scales (Blouin *et al. in review*; Funk *et al.* 2005; Funk *et al.* 2008; Monsen and Blouin 2003; Monsen and Blouin 2004; Hoffman and Blouin 2004a and 2004b). Thus, we assume that data obtained from each population sample are largely independent, even among neighboring populations.

For sample collection, DNA extraction, microsatellite amplification, and genotyping methods see Blouin *et al.*(in review) for *R. pretiosa*, Funk *et al.* (2005) for *R. luteiventris*, Monsen and Blouin (2003) for *R. cascadae*, and Hoffman *et al.* (2003) for *L. pipiens*. The microsatellite loci used for the four species are listed in Table A.2.

Estimation of effective population sizes

We obtained estimates of effective population size ($\widehat{N}_e$) for each population using the linkage disequilibrium (LD) method (Hill 1981) and the approximate Bayesian computation (ABC) method implemented in the program ONeSAMP (Tallmon *et al.* 2008). Two *R. pretiosa* populations were analyzed using Waples' (1989) temporal moment (TM) version of the temporal method (Nei and Tajima 1981). We used previously-published TM estimates for five of the *L. pipiens* populations (Hoffman *et al.* 2004). See Table 3.1 for the notation used in this paper.

The LD method is based on the principle that non-random associations between neutral alleles at different loci can be generated by genetic drift. In theory, the amount of linkage (i.e. gametic) disequilibrium in randomly-mating, isolated populations is entirely a function of drift and can be used to calculate $\widehat{N}_e$ (Hill 1981). This method provides an estimate of contemporary, local N_e in the previous generation, although LD generated over several generations can influence the estimate (Waples 2005). In addition to the assumptions of random mating and isolation, the LD method assumes selective neutrality of the genetic markers, no genetic substructure within the population, and non-overlapping generations. We calculated $\widehat{N}_e$ via the LD method using the program LDNe (Waples and Do 2008), which incorporates Waples' (2006) correction for the downward bias in $\widehat{N}_e$ that is introduced when the sample size is smaller than the true effective size. We ran LDNe under the random-mating model and report $\widehat{N}_{e(LD)}$ based on calculations which excluded rare alleles with frequencies

less than 0.02 when sample size (S) was greater than 25, following the recommendations of Waples and Do (2009). When $S \leq 25$, we adjusted the critical allele frequency (P_{crit}) to $1/2S < P_{crit} < 1/S$. Negative values for $\widehat{N}_e$ from the LD method are interpreted as infinity (Waples and Do 2009). We obtained confidence intervals using the jackknife option in LDNe, which performed better than the traditional parametric method in the simulation study of Waples and Do (2008).

ONeSAMP uses an ABC procedure to obtain $\widehat{N}_e$ by comparing 8 summary statistics calculated for each of 50,000 simulated populations to statistics from the real population under consideration (Tallmon *et al.* 2008). Each of the summary statistics (including a measure of linkage disequilibrium) is a function of N_e. ONeSAMP requires the specification of upper and lower bounds on the uniform prior distribution for N_e. For each population of the northwestern species we performed analyses under two different prior ranges: 2-200 and 2-2000. Priors were set at 2-2000 and 2-5000 for *L. pipiens* because $\widehat{N}_{e(LD)}$ and previously-published $\widehat{N}_e$ from the temporal method for some populations of this species were relatively large (Hoffman *et al.* 2004). The ONeSAMP input cannot include monomorphic loci or individuals missing data at more than one locus. For consistency, we used the same input data for both LDNe and the ONeSAMP.

For *R. pretiosa* populations RP1 and RP10 we collected samples in both 1999 (n = 28 for both RP1-A and RP10-A) and 2006 (n = 33 and 32 for RP1-B and RP10-B, respectively). This sampling allowed us to apply the TM method of N_e estimation for these populations. This method derives an estimate of N_e from the variance of neutral

allele frequencies between samples taken at two different times (Waples 1989). N_e estimated by the temporal method ($\widehat{N}_{e(TM)}$) applies to the time between the sampling events (Waples 2005). An estimate of the number of generations between the two temporal samples (*t*) is needed to apply the temporal method. Based on skeletochronology data from the RP10 population, males and females first breed at 2 and 3 yr, respectively (Blouin, unpublished data). We combined this information with data on age-specific mortality for *R. pretiosa* (Licht 1969; Licht 1974) to estimate the generation time in this species as approximately 3.1 yr. Thus, we assumed that $t = 2.25$ generations between the 1999 and 2006 samples. To estimate N_e with the temporal method (Nei and Tajima 1981), we used Waples' (1989) moment-based approach. We calculated $\widehat{N}_{e(TM)}$ using equations 9 and 12 from Waples (1989) and calculated confidence intervals using his equation 16. Because the temporal samples were taken non-destructively from adults, Waples' Plan I sampling design was appropriate (Waples 1989). This approach requires an estimate of the census size for the population at the time of the first sample. We calculated estimates separately using both $N = 500$ and $N = 2000$ because egg mass counts and mark-recapture work suggest that the RP1 and RP10 population sizes are each between 500 and 2000 individuals (K. McAllister and J. Bowerman, personal communication).

Genetic diversity

We calculated the average expected heterozygosity (H_e) and allelic richness (*AR*) for each population using the programs FSTAT (Goudet 1995) and

POPULATIONS (Langella 1999), respectively. We estimated *AR* via rarefaction at a common sample size of 15. To compare these measures of genetic diversity across populations and species, we excluded some populations due to small sample sizes and/or missing data. Thus, in our genetic diversity comparison we used only 20 *R. pretiosa* populations and 16 *R. cascadae* populations. For those populations that had more than one temporal sample (RP1, RP10, LP1-LP4), we used the more recent sample (e.g. RP1-B) in our comparisons.

Statistical analysis

We tested for differences in $\widehat{N}_e$, H_e, and *AR*, among the four species using the nonparametric Kruskal-Wallis test. When this test had a significant outcome, we conducted *post-hoc* analyses, using the Mann-Whitney U procedure to test for differences between each species pair. We tested for correlations between N_e estimates from the LD and ABC methods ($\widehat{N}_{e(LD)}$ and$\widehat{N}_{e(ABC)}$) within species. For these tests we excluded populations with negative $\widehat{N}_{e(LD)}$. To investigate relationships between geographic variables and $\widehat{N}_e$ within species, we tested for correlations between each of the genetic parameters ($\widehat{N}_e$, H_e, and *AR*) and each of three variables: latitude, longitude, and elevation. We log-transformed data, where appropriate, and tested the assumption of normality for variables using the Shapiro-Wilk test. Standard Pearson correlation (*r*) was used when both variables were normally-distributed; Spearman rank correlation (ρ) was used in all other cases. We used $\widehat{N}_{e(ABC,2000)}$ in all of these statistical tests. In cases where two of the geographic variables were significantly

correlated with each other and were each correlated with the same genetic parameter, we performed a linear regression analysis of residuals in an attempt to determine which geographic variable was driving the relationship. To do this, we regressed the genetic parameter on the first geographic variable and then used the residuals of this regression as the dependent variable of a second regression, using the second geographic variable as the independent variable in this case. We then repeated this procedure, switching the two geographic variables. If the regression of the residuals versus one of the two correlated geographic variables was significant but not the other, we took this as evidence that the first variable was driving the relationship.

To account for multiple comparisons, we adjusted the significance levels in the Mann-Whitney U correlation tests, and regression analyses using a Bonferroni correction. All of the statistical analyses were performed using R (R Core Development Team 2009).

Results

Single sample estimates of effective population size – LD and ABC methods

Estimates of N_e from the LD and ABC method are presented in Table A.3. $\widehat{N}_e$ from both methods was less than 50 for most populations of the northwestern frogs (Fig. 3.2; median $\widehat{N}_{e(LD)}$ for northwestern species = 31, excluding populations for which LDNe returned estimates of infinity; median $\widehat{N}_{e(ABC,2000)}$ for northwestern species = 27.78). Estimates for *L. pipiens* spanned a wider range and were mostly larger (median $\widehat{N}_{e(LD)}$ = 135; median $\widehat{N}_{e(ABC,5000)}$ = 120). No significant differences in

$\widehat{N}_e$ were found between any of the northwestern species, but $\widehat{N}_e$ is significantly larger in *L. pipiens* than in the other species (Table 3.2).

Effective population sizes within *R. pretiosa* and *R. cascadae* were not significantly correlated with latitude, longitude, or elevation (Table 3.3). However, $\widehat{N}_e$ is correlated with both longitude and elevation within *R. luteiventris* and *L. pipiens*. $\widehat{N}_e$decreases from west to east in *R. luteiventris* but increases in *L. pipiens* (Fig. 3.3). $\widehat{N}_e$decreases with elevation in both species and increases with latitude in *R. luteiventris*. In *R. luteiventris*, latitude and elevation were correlated and each was correlated with $\widehat{N}_e$ (Table 3.3). The regression analyses of residuals was unable to reveal which of these variables was driving the correlation with $\widehat{N}_e$ (i.e., neither regression was significant). In *L. pipiens*, results of the regression analyses suggest that, although longitude and altitude are correlated in this species, longitude appears to be the important variable with respect to $\widehat{N}_e$.The regression of the residuals from [$\widehat{N}_e$ versus longitude] versus elevation was not significant (R^2 = -0.03, p = 0.52), whereas the regression of the residuals from [$\widehat{N}_e$ versus elevation] versus longitude was significant (R^2 = 0.20, p = 0.02), although only before applying the Bonferroni correction.

LD method point estimates were infinity (i.e. had negative estimates) for 17 population samples; upper confidence limits were infinity for all but 27 samples. Estimates of infinity are returned when the signal in the genetic data can be attributed entirely to sampling error, rather than genetic drift, which is the case for a very large population or when the population sample contains too little information (Waples and

Do 2009). Point estimates or upper limits of infinity were never returned by the ABC method. However, $\widehat{N}_{e(ABC)}$ upper limits for a number *L. pipiens* populations were very large (in the many thousands), sometimes exceeding the upper prior (Table A.3).

Within the northwestern species, neither $\widehat{N}_{e(LD)}$ nor $\widehat{N}_{e(ABC)}$was consistently larger or smaller than the other. However, there was less variance in $\widehat{N}_{e(ABC)}$ among populations within species (Fig. 3.4). $\widehat{N}_{e(ABC)}$ tended to be larger than $\widehat{N}_{e(LD)}$ in *L. pipiens* (Fig. 3.4). Estimates from the two methods were significantly correlated only in *R. luteiventris* (ρ = 0.52, p <0.015) and *L. pipiens* (ρ = 0.73, p <0.0001). There was a positive, albeit nonsignificant, correlation between $\widehat{N}_{e(LD)}$ and $\widehat{N}_{e(ABC)}$ in *R. cascadae* (ρ = 0.296, p <0.299). Most of the point estimates and confidence limits from the ABC method increased slightly when the upper prior was increased (from 200 to 2000 for the northwestern species and from 2000 to 5000 for *L. pipiens*; Table A.3). However, ABC point estimates for 22 population samples decreased when the upper prior was increased. For most populations, the difference between the point estimates obtained under different upper priors was small.

Two-sample estimates of effective population size – TM method

The two $\widehat{N}_{e(TM)}$ values for *R. pretiosa* populations RP1 and RP10 were both between 60 and 70, regardless of whether the census size of the population at the time of the first sample was assumed to be 500 or 2000. With *N* set at 500, N_e was 65 for RP1 (95% CI: 21-1281) and 61 for RP10 (95% CI: 13-infinity). With *N* set at 2000, the respective N_e estimates were 70 (95% CI: 21-infinity) and 66 (95% CI: 13-

infinity). These point estimates from the temporal method are only slightly larger than estimates from the LD and ABC methods for the RP1 and RP10 populations (20-44 and 20-42, respectively).

Hoffman *et al.* (2004) previously estimated N_e for *L. pipiens* populations LP1-LP5 using the TM method (Waples 1989), the temporal method of Wang (2001), and the method of Wang and Whitlock (2003). Our single-sample estimates of N_e for these populations (from 2001) are mostly larger than the estimates of Hoffman *et al.* (2004; Table 3.4). Although there is a strong, positive correlation between $\widehat{N}_{e(ABC,2000)}$ and $\widehat{N}_{e(TM)}$ for these populations (r = 0.96, p = 0.008), other correlations comparing the single-sample estimates and the temporal estimates for these populations were non-significant, as were the correlations between single-sample estimates from the two time periods.

Genetic diversity

Expected heterozygosity (H_e) and allelic richness (*AR*) for each population are shown in Table A.1 and are plotted against each other in Fig. 3.5. The most notable pattern here is the much greater genetic diversity within many *L. pipiens* populations relative to the other three species. Results of the Kruskal-Wallis and Mann-Whitney U tests indicate that H_e and *AR* differ significantly between all species except between *R. luteiventris* and *R. cascadae* (Table 3.2). The *R. cascadae* populations generally have the highest levels of genetic diversity (H_e: 0.33-0.74; *AR*: 2.17-5.90) among the northwestern species. The populations of *R. pretiosa* tend to have very low levels of

genetic diversity (H_e: 0.14-0.50; *AR*: 1.64-3.98). The diversities of the *R. luteiventris* populations (H_e: 0.23-0.66; *AR*: 1.74-5.33) span the range of values seen in *R. cascadae* and *R. pretiosa*, although none are as low as the least genetically diverse *R. pretiosa* populations.

H_e and *AR* are highly correlated with each other and each is correlated with $\widehat{N}_e$ over all populations (Table 3.3). In *R. pretiosa*, none of the genetic parameters were correlated with latitude, longitude, or elevation. Genetic diversity is correlated with latitude in *R. luteiventris* and *R. cascadae,* although the patterns in the two species are reversed: H_e (and *AR*) increases at higher latitudes in *R. luteiventris* and decreases at higher latitudes in *R. cascadae*. As found by Funk *et al.* (2005), genetic diversity in *R. luteiventris* is also negatively correlated with elevation. There was a correlation between latitude and elevation for this species. The regression analyses of residuals for these two variables in *R. luteiventris* were both non-significant, preventing us from statistically disentangling their relationships with genetic diversity in this species. Longitude and elevation were correlated in *L. pipiens* and each is correlated with both H_e and *AR*. The regressions of the residuals from both [H_e versus longitude] versus elevation (R^2 = 0.02, p = 0.213) and [*AR* versus longitude] versus elevation (R^2 = -0.04, p = 0.981) were not significant. Conversely, regressions of the residuals from [H_e versus elevation] versus longitude (R^2 = 0.33, p = 0.0002) and [AR versus elevation] versus longitude (R^2 = 0.32, p = 0.0003) were significant. This suggests that longitude, rather than elevation, is the more important geographic variable with respect to genetic diversity in *L. pipiens*.

Discussion

The most striking finding of our study is that the three northwestern ranid frog species (and the Western populations of *L. pipiens*) appear to have very small contemporary effective population sizes (<50) and correspondingly low genetic diversities. Based on our results, a good estimate for the effective size of a northwestern ranid frog population would be about 20 or 30 individuals (Fig. 3.2). However, it is worth considering violations of assumptions of the methods and caveats about precision that affect how much faith to put in these numbers (Luikart et al., 2010; see Waples and Do, 2009 for in-depth discussions).

The ABC method is expected to suffer from less bias and imprecision than the LD method, given that $\widehat{N}_{e(ABC)}$ is based on linkage disequilibrium plus seven other parameters that are related to N_e (Tallmon *et al.* 2008; Luikart *et al.* 2010). $\widehat{N}_{e(ABC)}$ estimates were less variable among populations within species than $\widehat{N}_{e(LD)}$ estimates, even when negative estimates from the LD method were excluded. $\widehat{N}_{e(ABC)}$ estimates also had smaller nominal confidence intervals. Although the bias and precision of the ABC method under different biological situations has yet to be evaluated, our results suggest that the method probably is more reliable than the LD method. Nevertheless, estimates from the two methods were positively correlated in three species (significantly so in two), so even the simple LD method seems to be capturing much of the same information available from the more sophisticated method.

Samples from ranid frog populations are almost always in Hardy-Weinberg equilibrium (as are ours), so non-random mating or cryptic subdivision should not contribute to LD. Populations of the species in this study do violate the assumptions of discrete generations and (in some cases) closed populations. Given overlapping generations, the LD method actually estimates the effective number of breeding individuals (N_b) that produced the sampled cohort(s), which may not be the true, per-generation N_e. How single-sample estimates of N_e and N_b are related in age-structured populations is still unclear (Waples 2010). A small rate of immigration can cause mixture disequilibrium (downward biased estimates of N_e), although this effect is thought to be small at equilibrium (Waples and Do 2010). Even with gene flow as high as 10%, the LD method should provide estimates of local N_e (Waples 2010). Extremely high gene flow (e.g., >10%) could cause the estimate from each local population to approximate the value for the metapopulation. However, this will not be the case for most amphibians, which have highly structured metapopulations. Different estimators of N_e in these frogs (LD, ABC, temporal) gave very similar results, and our estimates are similar to those from other published studies of N_e in ranid frogs (Zeisset and Beebee 2003; Brede and Beebee 2006; Schmeller and Merila 2007; Ficetola *et al.* 2009; Phillipsen *et al*. 2009). Thus, there is no evidence to suggest that our single-sample estimates are strongly biased upwards or downwards. Our estimates should also be reasonably precise. Even for the sample sizes used in this study, the LD method is still considered reliable when true N_e is small (<50) (Waples and Do, 2009). Furthermore, if our estimates were highly imprecise, it seems unlikely

that our estimates would be as consistent among populations as they are (Fig 3.2). For these reasons, we believe it is reasonable to conclude that typical local effective sizes for populations of ranid frogs in western North America are in the range of a few tens of individuals.

Another striking result is that *L. pipiens* populations show a much wider range of effective sizes than the three northwestern species (Fig. 3.2), with very large populations in the East and small ones in the West. Why might *L. pipiens* be so different? One possibility is that because the northwestern species diverged relatively recently from a common ancestor (Hillis and Wilcox 2005), they share some characteristics (e.g. breeding behaviors) that predispose them to small N_e. Another possibility is that something about the habitat in the Northwest causes low effective sizes. That western populations of *L. pipiens* have effective sizes similar to those of the three northwestern *Rana* species is consistent with a habitat effect. The montane wetland habitats occupied by the northwestern *Rana* species exist as small patches surrounded by a matrix of rugged landscape. Census population size (and thus N_e) is likely to be restricted in small habitat patches, while inhospitable terrain between patches limits gene flow and can reduce N_e in isolated populations by preventing the introduction of new alleles through immigration. Evidence of gene flow limitation among populations of montane amphibians has been found in several studies (Monsen and Blouin 2004; Funk *et al.* 2005; Spear *et al.* 2005; Giordano *et al.* 2007; Koscinski *et al.* 2009). Landscape features that have been associated with gene flow restriction

include high ridges (Funk *et al.* 2005) and elevational differences between populations (Funk *et al.* 2005; Spear *et al.* 2005; Giordano *et al.* 2007).

Restricted gene flow coupled with small effective population sizes in montane habitats is a key feature of the Valley-Mountain Model of amphibian population structure proposed by Funk *et al.* (2005). This model was originally a generalization of the pattern of genetic structure found for the *R. luteiventris* populations included in the present study. By estimating N_e explicitly for these populations, we confirm that high-elevation populations of *R. luteiventris* do indeed have smaller $\widehat{N}_e$. Also, for *R. pretiosa*, we found that genetic diversity was highest in two of the lower-elevation, valley populations, RP1 and RP4. The RP4 population also had the highest $\widehat{N}_e$ in *R. pretiosa*, and egg mass count survey data suggest that it has one of the largest census sizes (M. Hayes, personal communication). In contrast to the northwestern species, *L. pipiens* generally occupies lowland valley habitats. If *L. pipiens* populations in these habitats maintain larger census sizes and experience high levels of gene flow they may tend to maintain larger effective sizes.

Patterns in N_e *among populations within each species*

Rana pretiosa. The low genetic diversity and small $\widehat{N}_e$ found in *R. pretiosa* highlight its status as an imperiled species. Genetic connectivity among populations of this species is also very low (Blouin *et al.* in review). It is possible that recent habitat fragmentation has decreased genetic connectivity among *R. pretiosa* populations resulting in reduced N_e and genetic diversity. Alternatively, small N_e may be a natural

characteristic of this species, in which case genetic diversity might have been lost from these populations over thousands of years.

Rana luteiventris. Populations of this species located at higher elevations (>1000 m) have smaller N_e than those found in valley habitats. In Funk *et al.* (2005), lower genetic diversities were reported for montane populations, but N_e was not estimated explicitly. Here we show that N_e is smaller for montane populations, suggesting that low genetic diversity is due in part to small N_e, not simply restricted gene flow among populations. In another study of *R. luteiventris*, Davis and Verrell (2005) estimated N_e directly for four populations using demographic data and found values (3.2-37.8) in the range of what we report here. Latitude and longitude are also correlated with the genetic parameters in *R. luteiventris* (Table 3.3), suggesting that genetic signatures of historical demographic processes are present in our data. However, latitude is also correlated with elevation for this species. Thus it is difficult to determine which of these two factors is more important to genetic diversity in *R. luteiventris*.

Rana cascadae. While most populations of this species have small N_e, they have generally higher genetic diversity than populations of the other two northwestern frogs. Given that *R. cascadae* and *R. pretiosa* are co-distributed in the Cascades Mountains and they have similarly small N_e estimates, the higher levels of genetic diversity seen in *R. cascadae* may reflect a greater level of connectivity among *R. cascadae* populations. R. *cascadae* populations are more numerous across the Cascades Mountains than those of *R. pretiosa* and there is a strong pattern of isolation

by distance among *R. cascadae* populations (Monsen and Blouin 2004), indicating stepping-stone gene flow among populations. Movements of up to 5.2 km between habitat patches have been documented for *R. cascadae* (Garwood and Welsh 2007) and the potential for gene flow among populations at a regional scale may be relatively high (Brown 1997). H_e in *R. cascadae* decreases with increasing latitude. The lower H_e of northern *R. cascadae* populations may indicate decreased connectivity at the periphery of the species range.

Lithobates pipiens. There is a remarkable west-to-east pattern of increasing N_e and genetic diversity in *L. pipiens*. A similar pattern was found previously with independent data from mitochondrial DNA: western populations have lower haplotype and nucleotide diversities than eastern populations (Hoffman and Blouin 2004a; Wilson *et al.* 2008). The largest $\widehat{N}_e$ for *L. pipiens* were found in populations of the Midwest and eastern regions identified by Hoffman and Blouin (2004a) as possible glacial refugia during the Pleistocene. Long-term environmental stability in these regions may have allowed large populations to persist for thousands of years. In contrast, N_e of the western populations has likely been influenced by extreme population bottlenecks and habitat fragmentation during the glacial cycles of the Pleistocene and during the drying/warming of western North America in the Holocene (Thompson *et al.* 1993). Moreover, the westernmost populations sampled in this study are located at the species' historical range margin. The small N_e and low diversities of these populations may also reflect their peripheral positions (Hoffman and Blouin 2004b).

Our single-sample estimates for the five *L. pipiens* populations with two temporal samples were only roughly similar in scale to the TM method estimates obtained by Hoffman *et al.* (2004). Although there is a correlation between $\widehat{N}_{e(ABC,2000)}$ and $\widehat{N}_{e(TM)}$, the latter are mostly smaller than the former. This could be because $\widehat{N}_{e(TM)}$ applies to a longer timeframe of 11-15 generations (Hoffman *et al.*, 2004) whereas $\widehat{N}_{e(ABC,2000)}$ applies to only a few generations. Population size fluctuations over the greater number of generations reflected by $\widehat{N}_{e(TM)}$ may have reduced this estimate relative to $\widehat{N}_{e(ABC,2000)}$.

How do our estimates of N_e *compare to those of other ranid frog species?*

In a summary of amphibian N_e estimates by Schmeller and Merila (2007), most of the estimates were less than 100, suggesting that amphibians tend to have small effective population sizes. Several recent studies had similar results (BeeBee 2009; Wang 2009; Mullen *et al.* 2010). Estimates of N_e based on microsatellite markers have been obtained for several ranid frog species. Estimates for three European species — *R. ridibunda* (2 populations; Zeisset and Beebee 2003), *R. temporaria* (2 populations, Brede and Beebee 2006; 2 populations, Schmeller and Merila 2007) and *R. latastei* (8 populations, Ficetola *et al.* 2009) — were in the range of sizes we found for our northwestern species. The values of $\widehat{N}_e$ we found for the eastern populations of *L. pipiens* are the largest reported for a ranid frog and are much higher than most estimates for amphibians (but see Funk *et al.* 2009). Based on our results and those of previous studies, it now seems reasonable to conclude that typical effective sizes for

populations of most ranid frogs in western North America are in the tens, rather than in in the hundreds or thousands. It will be interesting to explore what it is about the life history or habitat of *L. pipiens* in Eastern North America that make it an exception.

Summary

Most empirical studies of effective population size based on molecular genetic data have reported estimates of N_e (or N_b) from only a small number of populations per species, and few comparisons have been made among species (e.g. Jehle *et al*. 2001; Brede and Beebee 2006; Fraser *et al.* 2007). In this study, we estimated N_e for 90 populations across four frog species. The three northwestern species appear to have very small N_e, fitting the general pattern found in most previous studies on amphibians. Some populations of the fourth species, *L. pipiens*, have considerably larger N_e. There are significant differences in N_e and genetic diversity among species, and geographic trends in N_e among populations within species. In particular, there is a strong east-to-west trend of decreasing N_e in *L. pipiens*. Measures of genetic diversity (H_e and *AR*) are highly correlated with N_e estimates among populations, suggesting that variation in genetic diversity largely reflects variation in local N_e in this taxon. We discovered these patterns by using a comparative approach, analyzing data from a large number of populations within and among species. More studies like this might be useful for revealing the most important intrinsic (e.g. shared biological characteristics, shared evolutionary histories of populations in the same geographic

region) and extrinsic (e.g. landscape influences such as elevation and topography) factors that control N_e in different taxa.

Acknowledgements

We thank David Tallmon for assistance with the program ONeSAMP. Mark Christie provided helpful comments on an early draft of the manuscript. We thank the Oregon State University Center for Genome Research and Biocomputing (CGRB) for assistance with microsatellite genotyping.

Literature Cited

Beebee TJC (2009) A comparison of single-sample effective size estimators using empirical toad (*Bufo calamita*) population data: genetic compensation and population size-genetic diversity correlations. *Molecular Ecology* **18**, 4790-4797.

Blouin MS (2000) *Genetic data for recovery of the Oregon Spotted Frog,* Rana pretiosa*, in the Western United States, Final report for USGS FY 2000 Species at Risk Program project.*

Blouin, MS, Phillipsen IC, Monsen KJ (2010) Population structure and conservation genetics of the Oregon spotted frog, *Rana pretiosa*. *Conserv Genet.:* In press.

Brede E, Beebee TJC (2006) Large variations in the ratio of effective breeding and census population sizes between two species of pond-breeding anurans. *Biological Journal of the Linnean Society* **89**, 365-372.

Brown C (1997) *Habitat structure and occupancy patterns of the montane frog, Rana cascadae, in the Cascades Range, Oregon, at multiple scales: Implications for population dynamics in patchy landscapes*, Oregon State University Editor.

Charlesworth B (2009) Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* **10**, 195-205.

Collins JT, Taggart TW (2009) Standard common and current scientific names for North American amphibians, turtles, reptiles, and crocodilians. Sixth Edition. The Center for North American Herpetology.

Davis AB, Verrell PA (2005) Demography and reproductive ecology of the Columbia spotted frog (*Rana luteiventris*) across the Palouse. *Canadian Journal of Zoology* **83**, 702-711.

Easteal S (1985) The ecological genetics of introduced populations of the giant toad *Bufo marinus*. II. Effective population size. *Genetics* **110**, 107-22.

Federal Register (2009) Endangered and Threatened Wildlife and Plants; 90-Day Finding on a Petition to List the Northern Leopard Frog (*Lithobates* [=Rana] *pipiens*) in the Western United States as Threatened. *Federal Register* **74** (207), 55525-26.

Ficetola GF, Padoa-Schioppa E, Wang J, Garner TWJ (2009) Polygyny, census and effective population size in the threatened frog, *Rana latastei*. *Animal Conservation* **9999**.

Frankham R (1996) Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology* **10**, 1500-1508.

Fraser DJ, Hansen MM, Østergaard S, *et al.*. (2007) Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Molecular Ecology* **16**, 3866-3889.

Funk WC, Cannatella DC, Ryan MJ (2009) Genetic divergence is more tightly related to call variation than landscape features in the Amazonian frogs *Physalaemus petersi* and *P. freibergi*. *Journal of Evolutionary Biology* **22**, 1839-1853.

Funk WC, Pearl CA, Draheim HM, *et al.*. (2008) Range-wide phylogeographic analysis of the spotted frog complex (*Rana luteiventris* and *Rana pretiosa*) in northwestern North America. *Molecular Phylogenetics and Evolution* **49**, 198-210.

Funk WC, Blouin MS, Corn P, *et al.*. (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology* **14**, 483-96.

Funk WC, Tallmon D, Allendorf F (1999) Small effective population size in the long-toed salamander. *Mol Ecol* **8**, 1633-40.

Garwood J, Welsh H (2007) *Ecology of the Cascades Frog (*Rana cascadae*) and Interactions with Garter Snakes and Nonnative Trout in the Trinity Alps Wilderness, California* CA Department of Fish and Game.

Giordano AR, Ridenhour BJ, Storfer A (2007) The influence of altitude and topography on genetic structure in the long-toed salamander (*Ambystoma macrodactulym*). *Molecular Ecology* **16**, 1625-1637.

Goudet J (1995) FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *J Hered* **86**, 485-486.

Herreid II CF, Kinney S (1966) Survival of Alaskan Woodfrog (*Rana Sylvatica*) Larvae. *Ecology* **47**, 1039-1041.

Hill W (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetical Research*, 209-216.

Hillis DM, Wilcox TP (2005) Phylogeny of the New World true frogs (*Rana*). *Molecular Phylogenetics and Evolution* **34**, 299-314.

Hoffman E, Blouin M (2004a) Evolutionary history of the northern leopard frog: reconstruction of phylogeny, phylogeography, and historical changes in population demography from mitochondrial DNA. *Evolution* **58**, 145-59.

Hoffman E, Blouin MS (2004b) Historical data refute recent range contraction as cause of low genetic diversity in isolated frog populations. *Mol Ecol* **13**, 271-6.

Hoffman E, Schueler F, Jones A, Blouin MS (2006) An analysis of selection on a colour polymorphism in the northern leopard frog. *Molecular Ecology* **15**, 2627-2641.

Hoffman EA, Ardren WR, Blouin MS (2003) Nine polymorphic microsatellite loci for the northern leopard frog (*Rana pipiens*). *Molecular Ecology Notes* **3**, 115-116.

Hoffman EA, Schueler FW, Blouin MS (2004) Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern leopard frog. *Evolution* **58**, 2536-2545.

Jehle R, Arntzen J, Burke T, Krupa A, Hodl W (2001) The annual number of breeding adults and the effective population size of syntopic newts (*Triturus cristatus*, *T. marmoratus*). *Mol Ecol* **10**, 839-50.

Jones L, Leonard W, Olson D (2005) *Amphibians of the Pacific Northwest* Seattle Audubon Society, Seattle, WA.

Koscinski D, Yates AG, Handford P, Lougheed SC (2009) Effects of landscape and history on diversification of a montane, stream-breeding amphibian. *Journal of Biogeography* **36**, 255-265.

Langella O (1999) Populations 1.2.30: a population genetic software. CNRS UPR9034. Available at http://bioinformatics.org/~tryphon/populations/

Licht L (1969) Comparative breeding behavior of the red-legged frog (*Rana aurora aurora*) and the western spotted frog (*Rana pretiosa pretiosa*) in southwestern British Columbia. *Canadian Journal of Zoology* **47**, 1287-1299.

Licht L (1974) Survival of embryos, tadpoles, and adults of the frogs *Rana aurora aurora* and *Rana pretiosa pretiosa* sympatric in southwestern British Columbia. *Can J Zool* **52**, 613-27.

Luikart G, Ryman N, Tallmon D, Schwartz M, Allendorf F (2010) Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics* **11**, 355-373.

Monsen K, Blouin MS (2003) Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. *Molecular Ecology* **12**, 3275-86.

Monsen K, Blouin MS (2004) Extreme isolation by distance in a montane frog *Rana cascadae*. *Conservation Genetics* **5**, 827-835.

Mullen LB, Woods HA, Schwartz M, Sepulveda AJ, Lowe W (2010) Scale-dependent genetic structure of the Idaho giant salamander (*Dicamptodon aterrimus*) in stream networks. *Molecular Ecology* **19**, 898-909.

Nei M, Tajima F (1981) Genetic drift and estimation of effective populatoin size. *Genetics* **98**, 625-640.

Phillipsen IC, Bowerman J, Blouin MS (2009) Effective number of breeding adults in Oregon spotted frogs (*Rana pretiosa*): genetic estimates at two life stages. *Conservation Genetics*. **11**, 737-745.

R Development Core Team (2009) *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria.

Rorabaugh, JC (2005) *Rana pipiens* Schreber, 1782, Northern leopard frog. In: *Amphibian Declines: The Conservation Status of United States Species* (ed. Lannoo MJ), pp. 570-577. University of California Press, Berkeley.

Schmeller D, Merila J (2007) Demographic and Genetic Estimates of Effective Population and Breeding Size in the Amphibian *Rana temporaria*. *Conservation Biology* **21**, 142-151.

Seburn D, Seburn C (2000) Conservation priorities for the amphibians and reptiles of Canada. Prepared for World Wildlife Fund Canada and the Canadian Amphibian and Reptile Conservation Network.

Soule M (1976) Allozyme variation, its determinants in space and time. In: *Molecular Evolution* (ed. Ayala F), pp. 46-59. Sinaur Associates, Sunderland, MA.

Spear S, Peterson C, Matocq M, Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology* **14**, 2553-64.

Stebbins R (2003) *A Field Guide to Western Reptiles and Amphibians* Houghton Mifflin, New York.

Tallmon D, Koyuk A, Luikart G, Beaumont M (2008) onesamp: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources* **8**, 299-301.

Thompson R, Whitlock K, Bartlein P, Harrison S, Spaulding W (1993) Climatic Changes in the Western United States since 18,000 yr B.P. In: *Global Climates since the Last Glacial Maximum* (eds. Wright H, Kutzbach J, Webb T, *et al.*). University of Minnesota Press, Minneapolis.

U.S. Fish and Wildlife Service (2004) Species assessment and listing priority assignment form, *Rana luteiventris,* Columbia spotted frog (Great Basin DPS).

U.S. Fish and Wildlife Service (2005) Species assessment and listing priority assessment form. Western Washington Fish and Wildlife Office, Lacey, Washington. Available at http://ecos.fws.gov/docs/candforms_pdf/r1/D02A_V01.pdf

Wang IJ (2009) Fine-scale population structure in a desert amphibian: landscape genetics of the black toad (*Bufo exsul*). *Molecular Ecology* **18**, 3847-3856.

Wang J (2001) A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genet Res* **78**, 243-57.

Wang J (2005) Estimation of effective population sizes from data on genetic markers. *Philos Trans R Soc Lond B Biol Sci* **360**, 1395-409.

Wang J, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* **163**, 429-446.

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**, 379-91.

Waples RS (2002) Definition and estimation of effective population size in the conservation of endangered species.. In: *Population Viability Analysis* (eds. Beissinger SR McCullough D), pp. 147-168. University of Chicago Press, Chicago.

Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology* **14**, 3335-3352.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics* **7**, 167-184.

Waples RS (2010) Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. *Molecular Ecology Resources* DOI 10.1111/j.1755-0998.2010.02876.x

Waples RS, Yokota M (2007) Temporal Estimates of Effective Population Size in Species With Overlapping Generations. *Genetics* **175**, 219-233.

Waples RS, Do C (2008) ldne: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources* **8**, 753-756.

Waples RS, Do C (2009) Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* **3**, 244-262.

Wilson G, Fulton TL, Kendell K, *et al.*. (2008) Genetic diversity and structure in Canadian northern leopard frog (*Rana pipiens*) populations: implications for reintroduction programs. *Canadian Journal of Zoology* **86**, 863-874.

Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114-138.

Zeisset I, Beebee TJC (2003) Population genetics of a successful invader: the marsh frog *Rana ridibunda* in Britain. *Molecular Ecology* **12**, 639-646.

Table 3.1. Notation used.

H_e	Expected heterozygosity
AR	Allelic richness
N_e	Effective population size
N	Census population size
$\widehat{N}_e$ $\tilde{N}_e$	Estimate of effective population size
LD	Linkage disequilibrium
ABC	Approximate Bayesian computation
TM	Temporal moments method of Waples (1989)
$\widehat{N}_{e(LD)}$ $\tilde{N}_{e(LD)}$	Estimate of effective population size from the linkage disequilibrium method
$\widehat{N}_{e(ABC,200)}$ $\tilde{N}_{e(ABC,200)}$	Estimate of effective population size from the approximate Bayesian computation method, with an upper prior of 200
$\widehat{N}_{e(ABC,2000)}$ $\tilde{N}_{e(ABC,2000)}$	Estimate of effective population size from the approximate Bayesian computation method, with an upper prior of 2000
$\widehat{N}_{e(ABC,5000)}$ $\tilde{N}_{e(ABC,5000)}$	Estimate of effective population size from the approximate Bayesian computation method, with an upper prior of 5000
$\widehat{N}_{e(TM)}$ $\tilde{N}_{e(TM)}$	Estimate of effective population size from the temporal moments method

Table 3.2. Results of statistical tests of differences between species in effective population size $\widehat{N}_e$ (from the ABC method with upper prior of 2000 for the northwestern species and 5000 for *L. pipiens*) and measures of genetic diversity (H_e and *AR*). Statistics and their associated *p*-values are given for the Kruskal-Wallis test of difference in the median values among all species and the Mann-Whitney U test of pairwise differences between each pair of species. For the latter, the test statistic (*U*) is given in the lower half of the matrix and the *p*-value is given in the upper half. Test statistics that are significant are in bold.

	Kruskal-Wallis		**Mann-Whitney U**				
	X^2	*p*		*R. pretiosa*	*R. luteiventris*	*R. cascadae*	*L. pipiens*
$\widehat{N}_e$	**33.5**	<0.0001	*R. pretiosa*	-	0.213	0.398	<0.0001
			R. luteiventris	203	-	0.058	<0.0001
			R. cascadae	168	174	-	<0.0001
			L. pipiens	**89**	**627**	**452**	-
H_e	**55.0**	<0.0001	*R. pretiosa*	-	0.0002	<0.0001	<0.0001
			R. luteiventris	**107**	-	0.047	<0.0001
			R. cascadae	**17**	131	-	<0.0001
			L. pipiens	**18**	**634**	**378**	-
AR	**53.6**	<0.0001	*R. pretiosa*	-	0.001	<0.0001	<0.0001
			R. luteiventris	**454**	-	0.051	<0.0001
			R. cascadae	**18**	132	-	<0.0001
			L. pipiens	**24**	**637**	**380**	-

Table 3.3. Results of tests for correlation between genetic parameters ($\widehat{N}_e$, H_e, and AR) and geographic factors (latitude, longitude, elevation) for populations of the four frog species. In cases where the data were normally-distributed, standard Pearson correlation was performed and the test statistic is r. In all other cases, Spearman-Rank correlation was performed and the statistic is ρ. Significant outcomes are marked with asterisks. For the within-species correlation tests, significance was determined using a Bonferroni-adjusted cutoff α of 0.0042 (0.05 divided by 12; 4 species x 3 geographic factors = 12).

Correlations across all populations

	ρ	*p*-value
H_e vs $\ln(\widehat{N}_{e(ABC)})$	0.701*	<0.0001
AR vs $\widehat{N}_{e(ABC)}$	0.742*	<0.0001
H_e vs. $\ln(AR)$	0.980*	<0.0001

Correlations within species

	Rana pretiosa		*Rana luteiventris*		*Rana cascadae*		*Lithobates pipiens*	
	ρ (or *r*)	*p*-value	ρ (or *r*)	*p*-value	ρ (or *r*)	*p*-value	ρ (or *r*)	*p*-value
$\widehat{N}_{e(ABC)}$ vs Lat	0.00	0.9721	0.55*	0.0042	0.19 (*r*)	0.4184	0.24	0.2167
$\widehat{N}_{e(ABC)}$ vs Lon	0.42	0.0642	-0.61*	0.0012	0.22	0.3438	0.85* (*r*)	<0.0001
$\widehat{N}_{e(ABC)}$ vs Elev	-0.08	0.7431	-0.57*	0.0025	0.14 (*r*)	0.5566	-0.74*	<0.0001
H_e vs Lat	-0.07	0.7724	0.81*	<0.0001	-0.65* (*r*)	0.0062	0.17	0.4192
H_e vs Lon	-0.06	0.7890	-0.44	0.0243	-0.57	0.0208	0.93*	<0.0001
H_e vs Elev	-0.32	0.1434	-0.83*	<0.0001	-0.2 (*r*)	0.4569	-0.77*	<0.0001
AR vs Lat	-0.16	0.5077	0.77*	<0.0001	-0.49 (*r*)	0.05371	0.17	0.4297
AR vs Lon	-0.11	0.6465	-0.55*	0.0038	-0.5	0.0482	0.92*	<0.0001
AR vs Elev	-0.37	0.0907	-0.77*	<0.0001	-0.11 (*r*)	0.6781	-0.76*	<0.0001
Elev vs. Lat	-0.55	0.0132	-0.78*	<0.0001	-0.02 (*r*)	0.7028	-0.59*	0.0039
Elev vs Lon	-0.12	0.6169	0.49	0.0112	0.34	0.1360	-0.81*	<0.0001

Table 3.4. Single-sample and temporal method estimates of N_e for five *L. pipiens* populations. The first two columns are estimates from the single-sample ABC method, with upper priors of 2000 and 5000, respectively. ABC estimates from only the more recent sample for each population are shown. The last two columns are estimates reported in Hoffman et al (2004), from the temporal moments method of Waples (1989; $\widehat{N}_{e(TM)}$) and the method of Wang (2001; $\widehat{N}_{e(W)}$).

Population	$\widehat{N}_{e(ABC,2000)}$	$\widehat{N}_{e(ABC,5000)}$	$\widehat{N}_{e(TM)}$	$\widehat{N}_{e(W)}$
LP1	1870 (641-9572)	3093 (860-25497)	588 (378-1355)	324 (230-488)
LP2	7737 (2154-64223)	10996 (2253-275552)	1820 (660-∞)	469 (313-786)
LP3	97 (52-304)	188 (91-684)	410 (222-940)	102 (71-152)
LP4	2032 (797-13683)	1745 (533-14550)	1019 (490-∞)	243 (165-395)
LP5	764 (393-2262)	7251 (1720-89068)	420 (245-837)	205 (150-295)

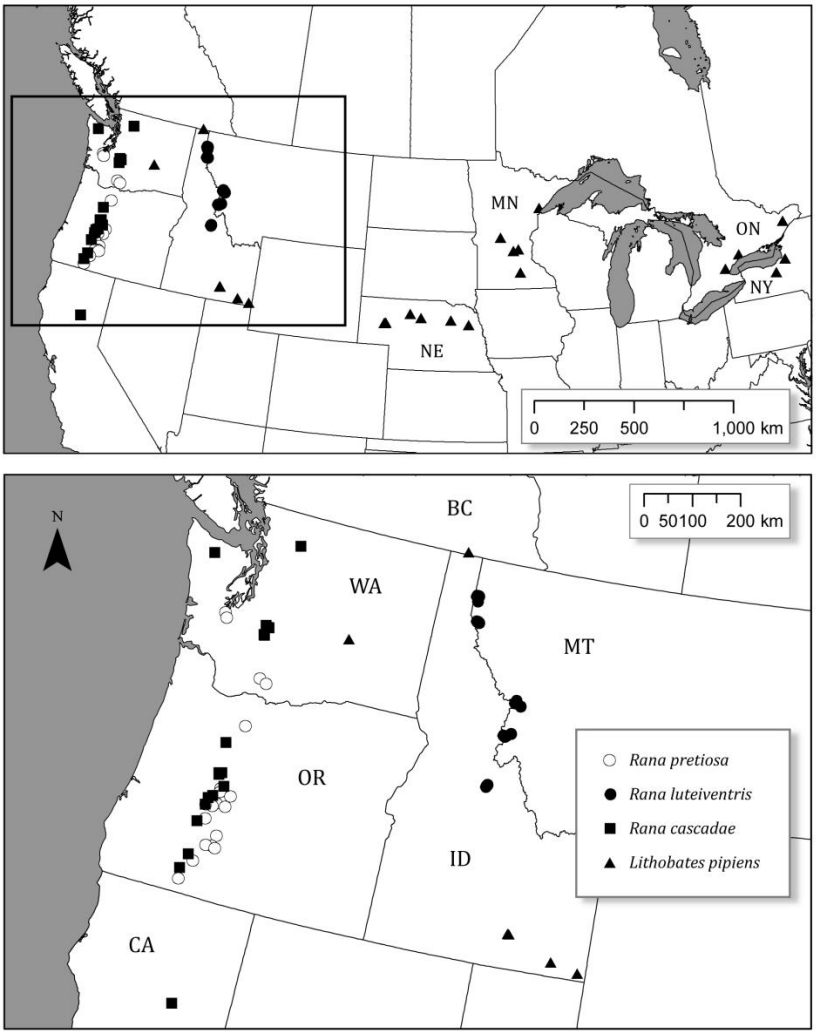

Figure 3.1. Collection localities for populations of the four frog species used in this study. The upper panel depicts the localities across North America. The inset black rectangle in the upper panel outlines the extent of the lower panel, which shows the northwestern localities in more detail. Abbreviations are shown for states/provinces in which samples were collected.

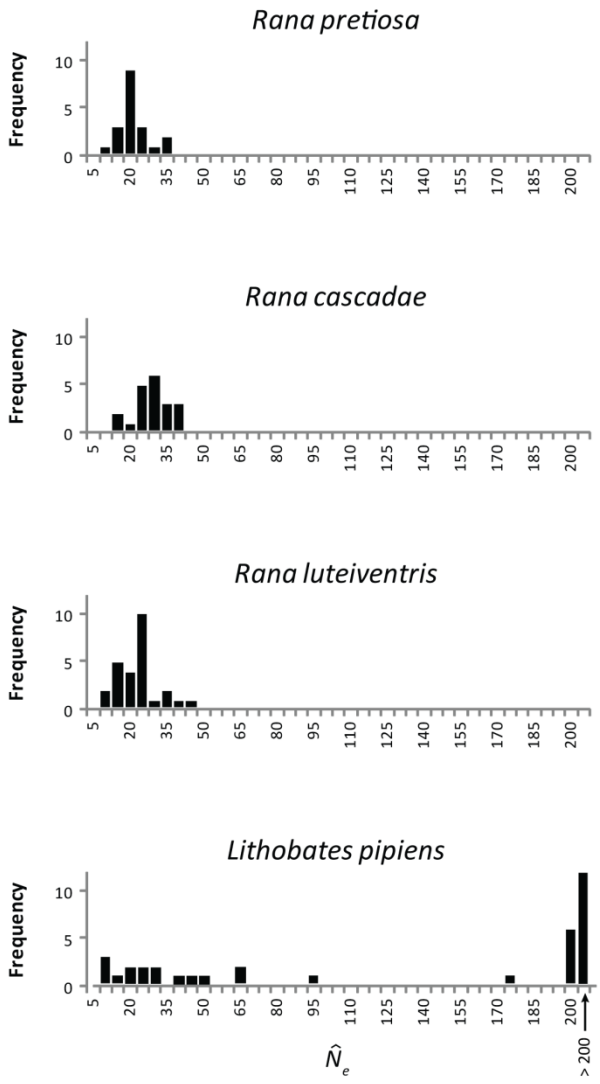

Figure 3.2. Estimates of effective population size ($\widehat{N}_e$) for the four frog species. Histograms of $\widehat{N}_e$ obtained using the ABC method with upper priors of 2000 are shown for populations of the northwestern species (*Rana pretiosa*, *R. cascadae*, and *R. luteiventris*) and estimates obtained with upper priors of 5000 are shown for *Lithobates pipiens*. There are no significant differences in $\widehat{N}_e$ among the northwestern species, for which all populations have $\widehat{N}_e$ < 50. *L. pipiens* has significantly higher median $\widehat{N}_e$ than any of the other species, with $\widehat{N}_e$ > 200 for some populations. See text for results of statistical tests.

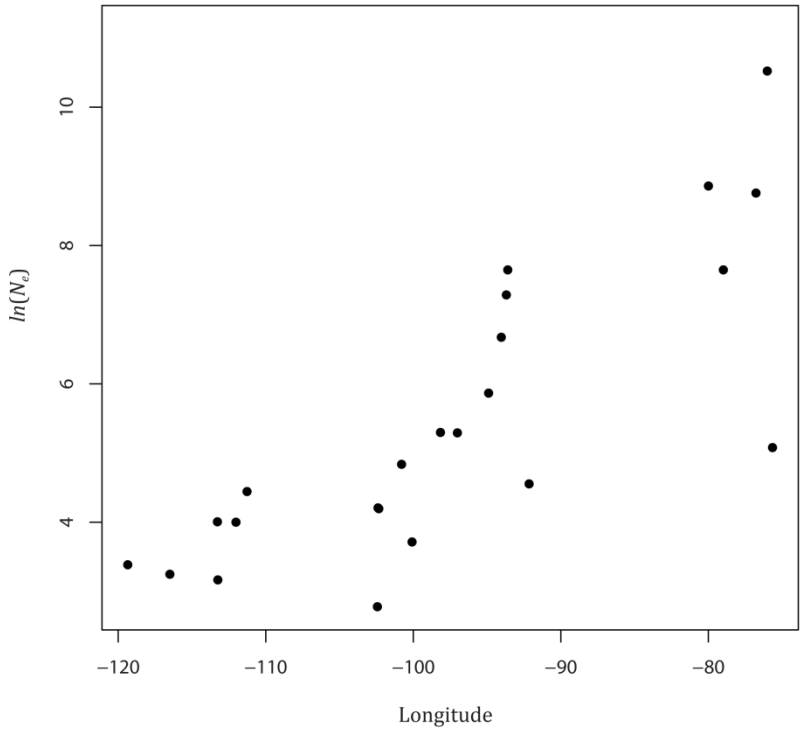

Figure 3.3. Plot of $\widehat{N}_e$ from the ABC method versus longitude for populations of *Lithobates pipiens*. A log transformation was applied to $\widehat{N}_e$ to improve the linear fit. There is a strong west-to-east pattern of increasing effective population size in this species. Measures of genetic diversity (H_e and *AR*) for these populations follow similar patterns.

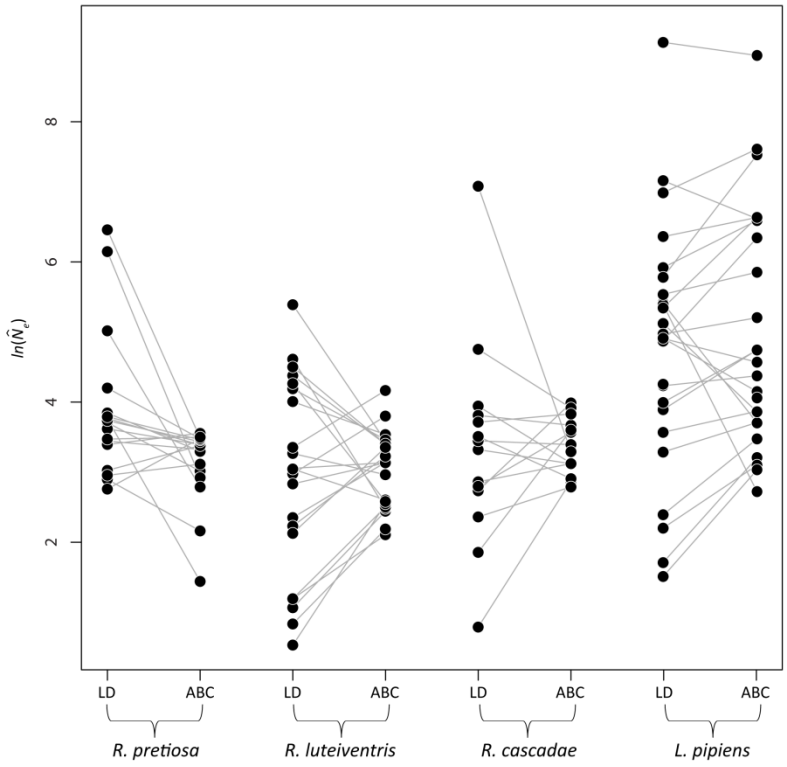

Figure 3.4. Plot of $\widehat{N}_e$ (on a logarithmic scale) from the LD and ABC methods, grouped by species. Gray lines between points connect estimates from the two methods for the same population. Populations for which $\widehat{N}_e$ from the LD method was negative (i.e. infinity) are not shown. For the northwestern (first three) species, the ABC estimates appear to have less variance within species, but do not have a consistent pattern of being greater than or less than the LD estimates. In *L. pipiens*, the ABC estimates have greater variance and are mostly larger than the LD estimates.

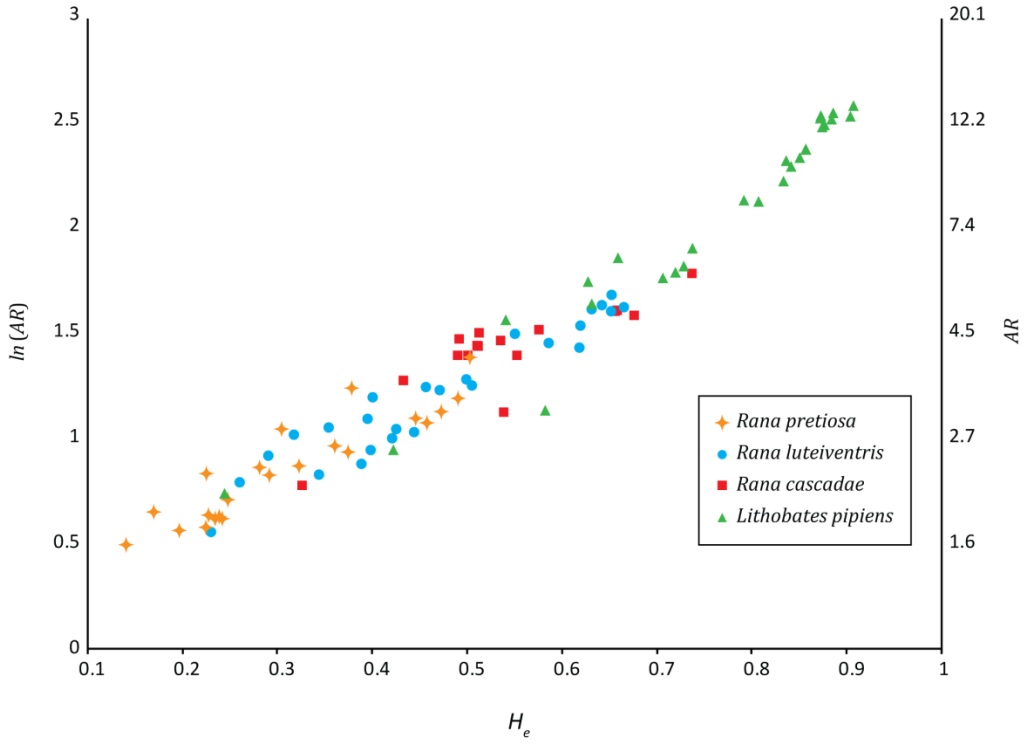

Figure 3.5. Plot of allelic richness (*AR*) versus expected heterozygosity (H_e) for populations of the four frog species. A log transformation was applied to *AR* to improve the linear fit. *Lithobates pipiens* populations have highest levels of genetic diversity. Among the other three species, *Rana cascadae* populations show the highest diversity. The populations with the lowest levels of diversity overall belong to *R. pretiosa*.

CHAPTER 4

A landscape genetics evaluation of connectivity among Cascades frog (*Rana cascadae*) populations in Olympic National Park

Ivan C. Phillipsen and Michael S. Blouin

Abstract

Genetic connectivity among animal populations in mountain landscapes may be attenuated by extreme topographic relief, and dispersal pathways between populations may more tortuous than direct due to differential permeability of landscape features. Previous landscape genetics studies of montane amphibians have identified aspects of topography, such as high ridges and elevation differences between populations as important determinants of population structure. We used a landscape genetics approach to infer the relative importance of landscape features in the genetic connectivity among 22 populations of the Cascades frog (*Rana cascadae*) in Olympic National Park, Washington. *R. cascadae* in the park have small effective population sizes (median $N_e = 31$) and showed high differentiation among populations (global $F_{ST} = 0.19$). There is a strong pattern of isolation-by-distance, with a large decrease in pairwise gene flow between populations separated by more than 10-15 km. To investigate how landscape controls gene flow, above and beyond the effects of distance *per se*, we constructed three alternative models of structural connectivity among populations and used a Random Forests statistical procedure to assess which of these explained the most variation in genetic distances among populations. For each connectivity model, 15 Landscape variables were measured along paths linking pairs of populations. We aimed to identify which of these variables are the most important predictors of genetic differentiation in *R. cascadae*. We evaluated these associations both within and between genetic clusters of populations, in order to determine how connectivity and the importance of landscape variables differ with spatial scale. We

identified six genetic groups using a Bayesian clustering analysis. The structural connectivity model that explained the most variation in genetic distance within these groups was one based on stream drainages. Forest cover and distance were the most important landscape variables in the stream model, impeding gene flow at the smaller scale. Between the genetic groups, the best structural connectivity model was one based on a minimum spanning tree, with subalpine pond habitats as nodes in the tree. Rock and ice cover acts as barriers and was the most important predictor in the between-group pond model. Distance along the paths was not the most important variable in the best-fitting models within or between genetic groups, suggesting that distance alone is not sufficient to explain the population structure of *R. cascadae*. Our results suggest that for species with relatively discrete habitats, such as pond-breeding amphibians, structural connectivity models based on graph theoretic networks (e.g. minimum spanning trees, relative neighborhood networks, etc.) may serve as relatively straightforward and parsimonious representations of structural connectivity.

Introduction

Studies in the burgeoning field of landscape genetics aim to reveal the influences of environmental factors on the distribution of genetic diversity within species (Manel et al, 2003). Knowledge of these influences is important for understanding microevolutionary processes in their spatial context and for the conservation management of species. Landscape genetics methods have been used to investigate a diversity of terrestrial and aquatic animals, including mammals (e.g.

Schwartz et al, 2009; Quéméré et al, 2010), birds (e.g. Pavlacky et al, 2009), reptiles (e.g. Clark et al, 2008), fish (e.g. Leclerc et al, 2008), and insects (e.g. Vandergast et al, 2007). Landscape genetics provides particularly useful tools for investigating the population genetics of amphibians. Frogs and salamanders tend to exhibit high levels of genetic differentiation among populations at small geographic scales (Chan and Zamudio, 2009). This differentiation is a consequence of generally small population sizes, natal philopatry, limited dispersal capabilities, and restricted habitat requirements (Waldman and McKinnon, 1993). Thus, compared to more vagile and generalist animals, amphibians are expected to have easily-detectable spatial genetic structure that is likely to have been influenced by landscape features.

Gene flow among populations of amphibians living in complex, rugged environments such as mountains may be especially limited and strongly associated with landscape features (Funk et al, 2005). In mountain landscapes, genetic connectivity may be attenuated by extreme topographic relief, and dispersal pathways between populations may be more tortuous than direct due to differential permeability of habitats. Previous landscape genetics studies of montane amphibians have identified aspects of topography, such as high ridges and elevation differences between populations as important determinants of population structure (Funk et al, 2005; Spear et al, 2005; Giordano et al, 2007). For some species, linear (i.e., Euclidean) distance between populations was found to be among the most important correlates of genetic differentiation (Spear and Storfer, 2008; Richards-Zawacki, 2009), while in others distance was a relatively poor predictor of differentiation (Murphy et al, 2010). Thus,

it cannot be assumed that linear distance alone is a suitable proxy for genetic distance in montane amphibians. Although few studies of amphibians have investigated the roles of landscape features at different geographic scales, there is evidence that some features may have a strong influence on gene flow at one scale but not another (Lee-Yaw et al, 2009; Koscinski et al, 2010; Murphy et al, 2010). More landscape genetics studies of montane amphibians are needed to identify general patterns in the way landscape features influence genetic diversity among species and how landscape influences differ across geographic scales and regions.

In this study, we used a landscape genetics approach to infer the relative importance of landscape features in the genetic connectivity among populations of the Cascades frog (*Rana cascadae*) in Olympic National Park (Washington State, USA). This species is restricted to subalpine wetlands in the mountains of the Pacific Northwest (Pearl and Adams, 2005). The pond, lake, and wet meadow habitats occupied by *R. cascadae* in the Olympic Mountains are shared by several other amphibian species (Corkran and Thoms, 2006): the long-toed salamander (*Ambystoma macrodactylum*), the northwestern salamander (*Ambystoma gracile*), the rough-skinned newt (*Taricha granulosa*), the Pacific treefrog (*Pseudacris regilla*), and the western toad (*Anaxyrus boreas*). Factors which influence genetic connectivity in *R. cascadae* may also be important for these amphibians, given their similar habitat requirements.

Genetic differentiation in *R. cascadae* among populations in the Cascades Mountains of Oregon and Washington fits a pattern of isolation by distance (IBD;

simple, linear distance), suggesting a 'stepping-stone' pattern of population structure (Monsen and Blouin, 2004). Migration between populations separated by more than 10 km was estimated as very low (Monsen and Blouin 2004). A 4-year mark-recapture study of *R. cascadae* in the Trinity Alps of California found that only 1% of marked individuals (19 of 1,955) moved between neighboring subalpine basins, which are separated by steep, dry ridges (Garwood, 2009). Thus, existing data suggest that gene flow in *R. cascadae* is able to maintain some connectivity at a broad scale (given the pattern of IBD in the Cascades Range) but is also strongly influenced by landscape features at a smaller geographic scale. Because *R. cascadae* is considered a declining species (Jennings and Hayes, 1994; Hammerson and Pearl, 2004), primarily due to population losses in the southern portion of its range (the causes are still uncertain; Fellers et al, 2008), there is a need for information on what influences connectivity among its populations.

We constructed three alternative models of connectivity among populations of *R. cascadae* in the Olympic Mountains and used a Random Forests statistical procedure (Breiman, 2001) to assess which of these explained the most variation in genetic distances among populations. For each structural connectivity model, 15 landscape variables were measured along paths linking pairs of populations. We wanted to identify which of these variables, given a particular connectivity model, are the most important predictors of genetic differentiation in *R. cascadae*. We evaluated these associations both within and between genetic clusters of populations, in order to

determine how connectivity and the importance of landscape variables differ with spatial scale.

Materials and Methods

Population Sampling

R. cascadae populations were sampled from 22 localities across the Olympic Mountains in Olympic National Park, WA in 2007 (Fig. 4.1; Table 4.1). These localities spanned the entire mountain range (~ 3,500 km^2). An average of 25 adult frogs was sampled opportunistically from pond and lake habitats at each locality. A total of 544 individuals were sampled. A single toe-clip tissue sample was taken from each individual and preserved in Drierite dessicant (W. A. Hammond Drierite Co., Xenia, OH).

Genotyping

Extractions of genomic DNA were performed using Qiagen DNeasy tissue kits (QIAGEN Inc.). Ten nuclear microsatellite loci were amplified using primers originally developed for *R. pretiosa* (Blouin et al, in review) and *R. muscosa* (Vredenburg et al, 2004). PCR was performed using Qiagen Multiplex PCR kits (QIAGEN Inc.). PCR conditions were as specified in the kit instructions: 30 s denaturation at 94° C, 90 s annealing at 60° C and 90 s extension at 72° C for 35 cycles, and a final 10 min extension at 72° C. Primer mixes were diluted in TE buffer. See Appendix B for primer sequences and multiplex primer mix volumes.

Each population sample was tested for conformity to Hardy-Weinberg equilibrium and linkage equilibrium with exact tests in GENEPOP (Raymond and Rousset, 1995). Genetic diversity for each population was quantified by expected heterozygosity (H_e) and allelic richness (*AR*), both averaged over all loci. *AR* was rarefied to a common sample size of 10. The two measures of genetic diversity were calculated using FSTAT (Goudet, 1995).

Population Structure

We quantified the overall level of population subdivision by calculating global F_{ST} (Weir and Cockerham, 1984) and its bootstrap confidence interval using FSTAT. We estimated genetic distances between each pair of populations with D_{ps}, the proportion of shared alleles (Bowcock et al, 1994), using the program Microsatellite Analyzer (Dieringer and Schlotterer, 2003). We also calculated pairwise F_{ST} and found that this measure was highly correlated with D_{ps}. We chose to carry out our statistical analyses using only D_{ps} because preliminary analysis showed that the landscape variables in our models could explain slightly more variation in D_{ps}. Exact tests of population differentiation were performed in GENEPOP, with 1000 batches and 1000 iterations per batch.

We evaluated the relationship of genetic distance to simple, Euclidean geographic distance in order to compare our results with those of the results of Monsen and Blouin (2004). We plotted pairwise $F_{ST}/(1- F_{ST})$ against pairwise geographic distance and tested for evidence of isolation by distance using the Mantel

test with 10,000 permutations. We also converted pairwise F_{ST} values into rough estimates of migration rate (i.e., the average number of migrants per generation) using the formula $Nm \approx (1- F_{ST})/4F_{ST}$ (Wright, 1968) and then plotted *Nm* against geographic distance.

To identify groups of genetically-similar populations, we used the individual-based Bayesian clustering program TESS (Chen et al, 2007). TESS incorporates spatial information from the sampling localities into Bayesian priors on the cluster membership of individuals via hidden Markov random fields (HMRF). Allele frequencies in the HMRF model are assumed to be most similar between neighboring localities and less similar between more distantly separated localities. We ran 20 independent TESS runs for each K_{max} (maximum number of population clusters) between 2 and 20. Parameters for each run consisted of 50,000 iterations (discarding the first 10,000 as burnin), no admixture, and an interaction parameter of 0.6. Following the TESS documentation guidelines, we determined the most likely number of clusters using the deviance information criterion (DIC) and post-processed the output using the program CLUMPP (Jacobsson and Rosenberg, 2007). For comparison, we also used the non-spatial clustering program STRUCTURE (Pritchard et al, 2000) to identify genetic groups. See Appendix B for STRUCTURE methods. As an alternative way of representing hierarchical population structure, we generated a neighbor-joining (NJ) tree of genetic distances (Nei's unbiased genetic distance; Nei, 1978) using the program POPULATIONS (Langella, 1999).

We estimated effective population size ($\widehat{N}_e$) for each population using the approximate Bayesian computation (ABC) method implemented in the program ONeSAMP (Tallmon et al. 2008). ONeSAMP generates $\widehat{N}_e$ by comparing 8 summary statistics calculated for each of 50,000 simulated populations to statistics from the real population under consideration (Tallmon et al. 2008). Each of the summary statistics is a function of N_e. Lower and upper bounds on the uniform prior distribution for N_e must be specified in ONeSAMP. We set these at 2 and 1000, respectively.

Landscape Genetics Analysis

To identify landscape features associated with population genetic structure in *R. cascadae*, we first constructed three alternative models of structural connectivity among *R. cascadae* populations, designated as 'Linear,' 'Pond,' and 'Stream' (Fig. 4.2). Each of these models represents a hypothesis about the paths that connect pairs of populations via gene flow. We used existing tools and custom Python scripts in the software ArcGIS 9.3 (ESRI) to create the connectivity models. Each model consisted of 231 pairwise paths (between all possible pairs of the 22 populations).

The Linear model was the simplest, with all population pairs connected by straight lines. The hypothesis represented by the Linear model was that gene flow between populations is direct and generally follows the shortest possible path. This model did not assume that connectivity is dependent upon specific landscape structures.

Pathways in the Pond and Stream models, on the other hand, were functions of clearly-defined structural characteristics of the landscape. The hypothesis underlying the Pond model was that gene flow between any two populations tends to occur in a "connect-the-dots" manner, i.e. through all intervening habitat patches along the network of patches. This pattern would be the case if dispersing frogs tend to move from pond to pond across the landscape. To construct the Pond model we generated a minimum spanning tree connecting all ponds, lakes, and wetlands located at elevations greater than 400 m. A minimum spanning tree is the shortest possible network (without loops) connecting point locations (i.e. habitat patches) distributed in two dimensions (Prim, 1957). Minimum spanning trees have been shown to be reliable representations of connectivity in ecological studies (Urban and Keitt, 2001; Vergara and Marquet, 2007). Paths between populations in the Pond model were constrained to follow branches of the minimum spanning tree.

In the Stream model, paths between populations followed stream and river drainages. The least cost paths analysis tool in ArcGIS 9.3 (ESRI, Redlands, CA) was used to generate the paths between each pair of populations. Given that Cascades frogs are not strictly aquatic and are capable of some overland movement (Garwood, 2009), the least cost paths were allowed to cross watershed boundaries. Map pixels designated as stream/river were assigned the minimal cost of movement (1) whereas non-stream/river pixels were assigned a much higher cost (1000). Although the cost values here are somewhat arbitrary, the large difference in values resulted in least cost paths with the desired characteristics of close conformation to stream/river drainages

and only rare deviations away from drainages. The National Hydrography dataset (United States Geological Survey) was used to construct the Pond and Stream models.

Next, we used the ArcGIS software to extract data on 15 landscape variables along each path connecting two populations (Table 4.2). These variables were extracted separately for each of the three structural connectivity models. We had *a priori* reasons to expect that these landscape variables might influence the movement of dispersing frogs and thus have an effect on the genetic differentiation of *R. cascadae* populations (Table 4.2). The landscape variables group into four categories: distance, topography, moisture/temperature, and cover. Explanations of the variables, their methods of calculation, and GIS data sources are listed in Table 4.2. Each pairwise path, originally in the form of a one-dimensional line, was first converted to a 150 m wide 'swath' so that total or mean values of landscape variables associated with the paths could be calculated from GIS data layers.

We evaluated the explanatory power of our connectivity models and the importance of the 15 landscape variables using the Random Forests procedure (Breiman, 2001). Murphy et al (2010) recently introduced the use of Random Forests in the context of landscape genetics. The method has great potential for applications in this field because it can deal with data that might be problematic for other statistical approaches, such as multiple linear regression. Random Forests is useful for making inferences and predictions from data that is noisy, autocorrelated, and non-independent (i.e. pairwise). Such data are common in landscape genetics studies. Furthermore, Random Forests can handle large numbers of predictor variables, is

insensitive to correlations among the predictor variables, and does not overfit the data (Breiman, 2001; Cutler et al, 2007). The procedure is an extension of classification and regression trees (De'ath and Fabricius, 2000), in which a dataset is recursively partitioned by predictor variables into groups that are as homogeneous as possible with respect to the response variable. In Random Forests, thousands of trees are built, each using a random two thirds of the observations and a random subset of the predictor variables. The fit of each tree is assessed via a cross-validation procedure where the remaining one third of the observations is run through the tree to generate predicted response values. The final output is a measure of the overall error rate (mean squared error) and percentage of variation explained (pseudo- r^2) by the Random Forests model, as well as an importance measure (I) for each of the predictor variables. The latter is a measure of how often a given predictor variable decreased the mean squared error of trees in the Random Forest model.

We used the Random Forests package in R (Liaw and Wiener, 2002; R Development Core Team, 2009) to perform these analyses. The response variable was genetic distance as measured by D_{ps} and the predictors were the 15 landscape variables. We built 5000 trees for each of our Random Forests models. The number of predictor variables randomly selected to build each tree was calculated as the total number of variables in the dataset divided by 3, as recommended by Liaw and Wiener (2002). We used the iterative variable selection procedure described by Murphy et al (2010), in which the importance value of each predictor variable is standardized by dividing it by the maximum importance possible in that model, resulting in a model

improvement ratio (MIR). Variables are then incrementally removed if their MIR falls below a cutoff importance value (0.1-1.0, iteratively increased by 0.1 increments). The final model selected was the one with the fewest number of predictor variables and the highest possible fit to the data.

It is possible that the model of structural connectivity that best explains genetic differentiation in *R. cascadae* at one spatial scale is not the best model at another scale. Likewise, the importance of landscape variables might differ across spatial scales (Lee-Yaw et al, 2009; Murphy et al, 2010). We distinguished spatial scales in our dataset by dividing the data for each of our three connectivity models into two subsets: within and between genetic groups. The genetic groups were those identified in our analysis of population structure using TESS (Fig. 4.1). Sample size was 40 for the 'within' dataset and 191 for the 'between' dataset. We performed a total of six independent Random Forests analyses, using within and between group datasets for each of the three connectivity models.

Results

Genetic diversity, N_e , *and population structure*

Over all populations there were no consistent deviations of the microsatellite loci from Hardy-Weinberg and linkage equilibrium. Genetic diversity was similar across populations, ranging from 0.64 to 0.80 for H_e and from 3.95 to 5.67 for *AR* (Table 4.1). Effective population size estimates were mostly small (median = 31;

range 14-154), which is consistent with results for other *R. cascadae* populations in the Cascades Range (Phillipsen, unpublished data).

The global F_{ST} was 0.119 (95% CI 0.103-0.133) and the average pairwise F_{ST} was 0.116 (range 0.004-0.242; Table B.2). These measures of genetic differentiation indicate a high level of population structuring in *R. cascadae* across the Olympic Mountains. In some cases, populations separated by as little as 3 km are significantly differentiated. The plot of pairwise $F_{ST}/(1- F_{ST})$ versus linear geographic distance indicates a strong pattern of isolation by distance (Fig 4.3). Because $\widehat{N}_e$ was similar across populations, pairwise *Nm* values derived from F_{ST} may largely reflect migration rates between populations, rather than genetic drift. Plotting *Nm* versus geographic distance revealed that gene flow appears to be highly restricted beyond pairwise distances of approximately 15 km (Fig 4.3). Similarly strong structure at this scale was found for *R. cascadae* populations in the Cascades Range (Monsen and Blouin 2003, 2004).

Results of the TESS analysis suggest that the 22 sampled populations cluster into six genetic groups (Fig 4.1). These groups were concordant with those in the NJ tree (Fig 4.4) and STRUCTURE results (Fig. B.1). The average pairwise geographic distances between populations within the genetic groups were 7.9, 16.9, and 10.8 km for the Linear, Pond, and Stream models, respectively. The corresponding distances between genetic groups were 39.8, 92.0, and 52.1 km.

Random Forests

Within the genetic groups, the structural connectivity model that explained the most variation in genetic distance was the Stream model (pseudo-r^2 = 0.55), followed by the Linear model (pseudo-r^2 = 0.42), and then the Pond model (pseudo-r^2 = 0.38; Table 4.3 and Fig. 4.5). The most important landscape variables in the Stream model were *forest* and *distance* (Fig 4.5). Partial dependency plots (Random forests output, not shown) indicated that genetic distance increased with increasing forest cover and distance. Between genetic groups, the best structural connectivity model was the Pond model (pseudo-r^2 = 0.74), followed closely by the Stream model (pseudo-r^2 = 0.72). Notably, the least supported model for the between group data was the Linear model (pseudo-r^2 = 0.63). *Rock-ice* was the most important variable in the between-group Pond model (i.e. this variable had the highest MIR value) and was the second most important variable in the other two connectivity models, after distance. Although *distance* is included in the Pond model it is not among the most important variables for that model (Fig. 4.5). Rather, *valley* and *ridge* were the next most important variables in the Pond model, after *rock-ice*. The partial dependency plots of *rock-ice*, *valley*, and *ridge* suggest that these features impede gene flow. *Distance* was not the most important variable in the best-fitting models within or between genetic groups, suggesting that distance alone is not sufficient to explain the population structure of *R. cascadae*.

Discussion

We found that populations of the Cascades frog in Olympic National Park show very strong genetic differentiation at a small spatial scale, even when compared to other temperate amphibians (Chan and Zamudio, 2009). This result is consistent with previous data on *R. cascadae* populations in the Cascades Range, mountains which are geographically isolated from the Olympic Mountains. To gain a deeper understanding of the processes underlying the population genetic structure we found, we used a landscape genetics approach to evaluate several population connectivity models and to highlight landscape variables that may influence gene flow in *R. cascadae*.

The best models of connectivity among populations were those based on structural characteristics of the landscape—ponds and streams— rather than simple linear connections. This difference is an important result for two reasons. First, it suggests that gene flow in *R. cascadae* occurs primarily along pathways defined by aquatic habitats. Although overland movements up to 5 km have been recorded in adult *R. cascadae* (Garwood, 2009), the majority of dispersing frogs are probably juveniles, which are less mobile than adults and are more prone to desiccation. Even if juvenile dispersal movements away from natal ponds are random (Semlitsch, 2008), the paths taken by successful dispersers (i.e. those that survive long enough to end up in a nonnatal breeding habitat) are likely to have followed aquatic habitats.

Second, this result demonstrates the value of evaluating ecologically realistic models of structural connectivity in addition to simpler models that assume no influence of landscape features. With minimal reliance on ‘expert opinion’ or other

subjective input, we used *a priori* information on habitat and dispersal behavior of *R. cascadae* to construct the Pond and Stream models. Similar findings that structural connectivity models based on landscape features are superior to simple linear models have been reported in previous landscape genetics studies of amphibians (Funk et al, 2005; Spear and Storfer, 2008; Richards-Zawacki, 2009; Spear and Storfer, 2010).

Within-group connectivity

Our data suggest that gene flow among *R. cascadae* populations at a small spatial scale, i.e., within genetically-similar groups, occurs via dispersal paths that follow stream drainages. Previous data support the hypothesis that streams facilitate connectivity in *R. cascadae*. A mark-recapture study of *R. cascadae* in California found that juvenile frogs were more likely to be found in streams than in any other habitat (Garwood, 2009). Juveniles were also the demographic stage most likely to disperse away from natal sites, a pattern common in anuran amphibians (Berven and Grudzien, 1990; Semlitsch, 2008). In the subalpine basins where *R. cascadae* live, pond and lake habitats are often interconnected by networks of streams (Naiman et al, 1992). An analysis of *R. cascadae* habitats in the Cascades Range of Oregon found that breeding habitats appear to have a greater number of stream connections than habitats where frogs do not breed (Brown, 1997). Connectivity based on streams in the Olympic Mountains was also shown to be important for a small mammal, the Pacific jumping mouse (*Zapus trinotatus*; Vignieri, 2005).

With respect to landscape variables retained in the Random Forests analysis of the within-group Stream model, the most important variable was *forest*. Forested areas may offer high resistance to dispersal because *R. cascadae* is adapted to habitats with minimal forest cover. In contrast, gene flow appears to be facilitated by forest cover in the coastal tailed frog (*Ascaphus truei*), a species adapted to densely-forested stream habitat in the Olympic Mountains (Spear and Storfer, 2008). Neither *rock-ice* nor *ridge* were retained for the Stream model. This may be because these cover types are much less extensive in the landscapes that separate populations belonging to the same genetic group than they are between populations in different genetic groups.

Distance was the second most important landscape variable in the Stream model and was included among the most important variables for most of the other Random Forests models. The inclusion of *distance* indicates that there is an isolation-by-distance effect both within and between genetic groups in *R. cascadae*, in addition to the effects of individual landscape features on gene flow.

Between-group connectivity

At the larger spatial scale that characterizes connectivity between genetic groups, dispersal pathways among populations were best approximated by a minimum spanning tree network, which connects ponds and lakes across the Olympic Mountains. Under the reasonable assumption that most of the subalpine wetland habitats in this region support *R. cascadae* populations, this result suggests that gene flow between genetic groups occurs in a stepping-stone fashion through occupied

habitat patches. The Pond model is a simple, easily-defined representation of the structural connectivity of *R. cascadae* habitats. More sophisticated connectivity models could be constructed by least cost path analyses of habitat distribution maps (e.g.,Wang and Summers, 2010). However, a greater number of assumptions and parameters are involved in this approach. Our results suggest that for species with relatively discrete habitats, such as pond-breeding amphibians, structural connectivity models based on graph theoretic networks (e.g. minimum spanning trees, relative neighborhood networks, etc.) may serve as relatively straightforward and parsimonious representations of structural connectivity.

The most important landscape variables in the between-group Pond model were associated with elevation (*valley* and *ridge*) and barren alpine terrain (*rock-ice*). Deep valleys separating *R. cascadae* habitats may function as barriers to gene flow between genetic groups. Likewise, high, rocky ridges and persistent ice fields may act as barriers at this scale. Ridges and elevation differences between populations are also important limiters of gene flow in the Columbia spotted frog (*Rana luiteventris*; Funk et al, 2005) and the long-toed salamander (*Ambystoma macrodactylum*; Giordano et al, 2007). Ridges in the Olympic Mountains are also associated with population structure in the Pacific jumping mouse (Vignieri, 2005). Interestingly, mark-recapture and radio-tracking data show that high ridges are sometimes traversed by adult *R. cascadae*. Adults are sometimes encountered some distance away from aquatic habitats (personal observation). Although adult frogs may be able to cross some topographic barriers, it is possible that most dispersing juveniles cannot. Most gene

flow probably occurs via the dispersal of juvenile frogs (which far outnumber and are less philopatric than adults). Thus, the landscape features that effect the movements of juveniles are what should influence population structure.

Note that for the between-group data, the Pond and Stream models were very similar in their ability to predict genetic distance. This may reflect the interconnected nature of these aquatic habitats in the Olympic Mountains. Our models are representations of dispersal pathways averaged over many generations. We cannot assume that dispersal always conforms to only one model. In reality, dispersal may sometimes occur via stream drainages or from pond to pond at other times, or even along straight line paths between populations. In other words, our models of structural connectivity are not necessarily mutually exclusive.

Landscape genetics of montane amphibians

This study provides further evidence that topographic features in mountain landscapes strongly influence the population structures of amphibians. High ridges have been identified as important barriers for three amphibians in western North America: the Columbia spotted frog, *Rana luteiventris* (Funk et al, 2005); the western toad, *Anaxyrus boreas* (Murphy et al, 2010); and the long-toed salamander, *Ambystoma macrodactylum* (Giordano et al, 2007). *R. luteiventris* is a close-relative of *R. cascadae* (Hillis and Wilcox, 2005) and the two species have similar autecologies (Nussbaum et al, 1983). Ridges are predictors of differentiation in *A. boreas* between genetic groups but not within them, similar to our results for *R. cascadae*. Populations

of *R. luteiventris* and *A. macrodactylum* in low elevation valleys are differentiated from populations located at high elevation. *R. cascadae* populations occur within a comparatively narrow elevation range, preventing their simple classification into high and low elevation groups. However, the retention of the *valley* and *ridge* variables in the best between-group Random Forests model (i.e. the Pond model) suggests that elevation also plays a significant role in the population structure of this species. If ridges and elevation differences generally act as barriers for amphibians, their effects may be most dramatic on species in tropical environments. Compared to temperate amphibians, tropical species may be less able to tolerate the cooler temperatures experienced on high mountain ridges (Janzen, 1967). Further landscape genetics studies of montane amphibians across a variety of environments might reveal that the effects of landscape features differ among regions.

Slope is another variable that was retained in models for *R. cascadae*, although it was not among the most important variables. Slope was identified as important for *Ascaphus truei* in the Olympic Mountains (Spear and Storfer, 2008) and *Atelopus varius* in Panama (Richards-Zawacki, 2009). Topographic features may not always be strongly associated with population structure in amphibians. Indeed, no support was found for the hypothesis that ridges influence gene flow in two montane frogs in Asia: *R. kukunoris* (Zhao et al, 2009) and *R. chensinensis* (Zhan et al, 2009).

Streams and rivers are very common features of mountain landscapes that may be important dispersal corridors for amphibians (Olsen et al, 2007). We found evidence that structural connectivity within genetically-similar groups of populations

of *R. cascadae* is associated with streams. Structural connectivity based on streams was also supported for *Ascaphus montanus* in the Rocky Mountains (Spear and Storfer, 2010) and there is suggestive evidence that gene flow in *Atelopus varius* is associated with streams (Richards-Zawacki, 2009).

Summary

The Cascades frog exhibits strong population structure, even at a small geographic scale (< 15 km), as revealed by analyses of microsatellite genetic variation among populations in the Cascades Range (Monsen and Blouin 2003, 2004) and Olympic Mountains. We used a recently-introduced landscape genetics method to demonstrate that, for populations in the Olympic Mountains, the simplest model of connectivity based on linear (Euclidean) paths does not explain the variation in genetic distances among populations as well as models based on the structural connectivity of streams and ponds. Our results suggest that gene flow in *R. cascadae* is restricted by forest cover at a small spatial scale (i.e., within genetic groups) and by barren alpine terrain, ridges, and valleys at larger spatial scale (i.e., between genetic groups). It is possible that these latter features are what define the boundaries between genetic groups.

Warming global temperatures may cause shifts in the altitudinal distributions of montane plant communities in the Pacific Northwest, so that the subalpine wetland habitats presently occupied by *R. cascadae* and several other amphibian species may be engulfed by (i.e., succeeded by) forests (Zolbrod and Peterson, 1999; Fagre et al,

2003). Thus, the habitats of these montane amphibians may become increasingly scarce, fragmented, and isolated by restricted gene flow in the coming decades. Melting glaciers and reduction of persistent snow fields may also alter the patterns of connectivity among populations (Fagre et al, 2003). This study thus provides a basis for forecasting the evolutionary responses of *R. cascadae* to global warming in the mountains of the Pacific Northwest.

Acknowledgements

We thank Jerry Freilich and Steve Fradkin of Olympic National Park for assistance with research permits and logistical help. Joshua Weber, Laura McMullen, Karen Kiemnec-Tyburczy, Sarah Eddy, Jeremy Noring, and Sam Hartsfield provided assistance with field sampling. Michael Adams and Deanna Olson provided helpful information on sampling localities and *R. cascadae* natural history. We thank the Oregon State University Center for Genome Research and Biocomputing (CGRB) for assistance with microsatellite genotyping. Funding from the Society of Wetland Scientists provided partial funding for this research.

Literature Cited

Berven K, Grudzien T (1990) Dispersal in the wood frog (*Rana sylvatica*): Implications for genetic population structure. *Evolution* **44**, 2047-2056.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, et al. (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**, 455-457.

Breiman L (2001) Random Forests. *Machine Learning* **45**, 5-32.

Brown C (1997) *Habitat structure and occupancy patterns of the montane frog,* Rana cascadae*, in the Cascades Range, Oregon, at multiple scales: Implications for population dynamics in patchy landscapes*, Oregon State University.

Chan LM, Zamudio KR (2009) Population differentiation of temperate amphibians in unpredictable environments. *Molecular Ecology* **18**, 3185-3200.

Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes* **7**, 747-756.

Clark RW, Brown WS, Stechert R, Zamudio KR (2008) Integrating individual behaviour and landscape genetics: the population structure of timber rattlesnake hibernacula. *Molecular Ecology* **17**, 719-730.

Corkran C, Thoms C (2006) *Amphibians of Oregon, Washington, and British Columbia* Lone Pine Publishing, Auburn, WA.

Cutler DR, Edwards TC, Beard KH, et al. (2007) Random Forests for classification in ecology. *Ecology* **88**, 2783-2792.

De'ath G, Fabricius KE (2000) Classification and regression tress: a powerful yet simple technique for ecological data analysis. *Ecology* **81**, 3178-3192.

Dieringer D, Schlotterer C (2003) microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* **3**, 167-169.

Evanno G, Regnauts S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Fagre DB, Peterson DL, Hessl AE (2003) Taking the Pulse of Mountains: Ecosystem Responses to Climatic Variability. *Climatic Change* **59**, 263-282.

Fellers GM, Pope K, stead JE, Koo M, Welsh H (2008) Turning population trend monitoring into active conservation: Can we save the Cascades frog (*Rana cascadae*) in the Lassen region of California? *Herpetological Conservation and Biology* **3**, 28-39.

Funk W, Blouin M, Corn P, et al. (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology* **14**, 483-96.

Garwood J (2009) *Spatial Ecology of the Cascades Frog: Identifying Dispersal, Migration, and Resource Uses at Multiple Spatial Scales*, Humbolt State University.

Giordano AR, Ridenhour BJ, Storfer A (2007) The influence of altitude and topography on genetic structure in the long-toed salamander (*Ambystoma macrodactulym*). *Molecular Ecology* **16**, 1625-1637.

Goudet J (1995) FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *J Hered* **86**, 485-486.

Hedrick P (2005) A standardized genetic differentiation measure. *Evolution* **59**, 1633-1638.

Hillis DM, Wilcox TP (2005) Phylogeny of the New World true frogs (Rana). *Molecular Phylogenetics and Evolution* **34**, 299-314.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806.

Janzen DH (1967) Why Mountain Passes are Higher in the Tropics. *The American Naturalist* **101**, 233.

Jennings M, Hayes M (1994) *Amphibian and Reptiles Species of Special Concern in California*.

Langella O (1999) Populations 1.2.30: a population genetic software. CNRS UPR9034. Available at http://bioinformatics.org/~tryphon/populations/

Leclerc E, Mailhot Y, Mingelbier M, Bernatchez L (2008) The landscape genetics of yellow perch (*Perca flavescens*) in a large fluvial ecosystem. *Molecular Ecology* **17**, 1702-1717.

Lee-Yaw JA, Davidson A, McRae BH, Green DM (2009) Do landscape processes predict phylogeographic patterns in the wood frog? *Molecular Ecology* **18**, 1863-1874.

Liaw A, Wiener M (2002) Classification and regression by Random-Forest. *R News* **2**, 18-22.

Manel S, Schwartz M, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution* **18**, 189-197.

Monsen K, Blouin M (2003) Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. *Molecular Ecology* **12**, 3275-86.

Monsen K, Blouin M (2004) Extreme isolation by distance in a montane frog Rana cascadae. *Conservation Genetics* **5**, 827-835.

Murphy MA, Evans JS, Storfer A (2010) Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology* **91**, 252-261.

Naiman RJ, Beechie TJ, Benda LE, et al. (1992) Watershed Management: Balancing Sustainability with Environmental Change. In: *Watershed Management: Balancing Sustainability with Environmental Change* (ed. Naiman RJ), pp. 127-188. Springer-Verlag, New York, USA.

Olson DH, Anderson PD, Frissell CA, Welsh Jr. HH, Bradford DF (2007) Biodiversity management approaches for stream-riparian areas: Perspectives for Pacific Northwest headwater forests, microclimates, and amphibians. *Forest Ecology and Management* **246**, 81-107.

Pavlacky Jr. D, Goldizen AW, Prentis PJ, Nicholls JA, Lowe AJ (2009) A landscape genetics approach for quantifying the relative influence of historic and contemporary habitat heterogeneity on the genetic connectivity of a rainforest bird. *Molecular Ecology* **18**, 2945-2960.

Pearl CA, Adams MJ (2005) Amphibian Declines: The Conservation Status of United States' Species. In: *Amphibian Declines: The Conservation Status of United States' Species* (ed. Lannoo M), pp. 538-540. University of California Press, Berkeley, CA.

Pike RJ, Wilson SE (1971) Elevation-Relief Ratio, Hypsometric Integral, and Geomorphic Area-Altitude Analysis. *Geological Society of America Bulletin* **82**, 1079-1084.

Prim R (1957) Shortest connection networks and some generalizations. *Bell Systems Technical Journal* **36**, 1389-1401.

Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-59.

Quéméré E, Crouau-Roy B, Rabarivola C, EEL JR, Chikhi L (2010) Landscape genetics of an endangered lemur (*Propithecus tattersalli*) within its entire fragmented range. *Molecular Ecology* **19**, 1606-1621.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *J Hered* **86**, 248-249.

R Development Core Team (2009) *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria.

Richards-Zawacki CL (2009) Effects of slope and riparian habitat connectivity on gene flow in an endangered Panamanian frog, *Atelopus varius*. *Diversity and Distributions* **15**, 796-806.

Sappington JM, Longshore KM, Thompson DB (2007) Quantifying Landscape Ruggedness for Animal Habitat Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert. *Journal of Wildlife Management* **71**, 1419-1426.

Schwartz MK, Copeland JP, Anderson NJ, et al. (2009) Wolverine gene flow across a narrow climatic niche. *Ecology* **90**, 3222-3232.

Semlitsch RD (2008) Differentiating Migration and Dispersal Processes for Pond-Breeding Amphibians. *Journal of Wildlife Management* **72**, 260-267.

Spear S, Peterson C, Matocq M, Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology* **14**, 2553-64.

Spear SF, Storfer A (2008) Landscape genetic structure of coastal tailed frogs (*Ascaphus truei*) in protected vs. managed forests. *Molecular Ecology* **17**, 4642-4656.

Urban D, Keitt T (2001) Landscape connectivity: a graph-theoretic perspective. *Ecology* **82**, 1205-1218.

Vandergast A, Bohonak A, Weissman D, Fisher R (2007) Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history: phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: *Stenopelmatus*). *Molecular Ecology* **16**, 977-992.

Vergara P, Marquet P (2007) On the seasonal effect of landscape structure on a bird species: the thorn-tailed rayadito in a relict forest in northern Chile. *Landscape Ecology* **22**, 1059-1071.

Vignieri S (2005) Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). *Molecular Ecology* **14**, 1925-37.

Vredenburg V, Tunstal T, Bingham R, et al. (2004) Patterns of habitat use and movement of Rana muscosa in the northern Sierra Nevada with comparisons to populations in the southern Sierra Nevada, with additional information on the biogeography of the species. CA Department of Fish and Game.

Waldman B, McKinnon JS (1993) The Natural History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives. In: *The Natural History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives*. University of Chicago Press, Chicago.

Wang IJ, Summers K (2010) Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology* **19**, 447-458.

Weir B, Cockerham C (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**, 1358-1370.

Wright S (1968) *Evolution and the genetics of populations.* University of Chicago Press, Chicago.

Zhan A, Li C, Fu J (2009) 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**, 17.

Zhao S, Dai Q, Fu J (2009) Do rivers function as genetic barriers for the plateau wood frog at high elevations? *Journal of Zoology* **279**, 270-276.

Zolbrod A, Peterson DL (1999) Response of high-elevation forests in the Olympic Mountains to climatic change. *Canadian Journal of Forest Research* **29**, 1966-1978.

Table 4.1. Population sampling localities for *Rana cascadae* in the Olympic Mountains.

Pop	*n*	UTM E	UTM N	Elev (m)	H_e	Mean # alleles	Mean $AR_{(8)}$
1	35	435127	5310413	955	0.67	5.9	4.3
2	10	437491	5310599	860	0.68	4.5	4.3
3	21	438668	5307935	1102	0.72	6.0	4.9
4	32	441654	5307719	1287	0.68	6.0	4.7
5	10	445588	5307181	1309	0.64	5.0	4.7
6	27	444765	5313686	1320	0.64	4.8	4.0
7	34	473853	5303659	1542	0.68	6.3	4.6
8	23	484218	5297662	1560	0.66	5.8	4.7
9	75	458657	5285745	1106	0.71	8.6	5.4
10	18	444857	5272583	981	0.71	7.0	5.5
11	20	445587	5272185	974	0.72	6.8	5.3
12	18	446633	5274734	1170	0.69	7.2	5.6
13	10	459758	5264644	1098	0.80	5.9	5.7
14	28	461345	5264155	1218	0.78	7.5	5.7
15	31	463382	5265956	1218	0.75	6.3	5.0
16	24	472607	5277324	1339	0.69	5.8	4.7
17	12	475332	5267523	1338	0.72	5.5	4.9
18	15	474332	5269043	1396	0.75	6.2	5.3
19	19	475731	5269162	1472	0.71	5.5	4.6
20	19	483775	5275365	1458	0.69	5.7	4.7
21	42	481624	5274149	1477	0.75	7.5	5.5
22	16	480815	5273087	1435	0.71	6.2	5.1

Table 4.2. Landscape variables used as predictors in Random Forests analyses. Variables are grouped into four categories: distance, topography, moisture/temperature, and cover. *A priori* expectations for the effects of each variable on gene flow are listed as hypothesized relationships. Calculations performed in ArcGIS 9.3 are given, along with GIS source data.

Landscape Variable	Hypothesized relationship to gene flow in *R. cascadae*	Calculation	Source*
Distance			
distance	Natal philopatry and/or inability to travel long distances may limit distances traveled by dispersing frogs	Total distance along path, adjusted for topography	DEM
Topography			
slope	Steep slopes may impede the movement of dispersing frogs.	Mean slope along path, in degrees	DEM
ruggedness	Rugged terrain may impede the movement of dispersing frogs.	Mean elevation ruggedness along path, calculated using a 350 X 350 m window (Sappington et al, 2007)	DEM
elevation relief ratio (ERR)	Terrain with large changes in elevation may impede the movement of dispersing frogs.	Mean elevation along path minus minimum elevation divided by relief (max elevation minus min elevation; Pike and Wilson, 1971)	DEM
ridge	Dispersing frogs may be intolerant of conditions at elevations much higher than those they typically occupy.	Total number of pixels on path multiplied by percentage of path covered. Areas with elevation > 2 standard deviations above the mean elevation for the sampled populations.	DEM
valley	Dispersing frogs may be intolerant of conditions at elevations much lower than those they typically occupy.	Total number of pixels on path multiplied by percentage of path covered. Areas with elevation < 2 standard deviations below the mean elevation for the sampled populations.	DEM
optimum elevation	Gene flow may be less restricted at elevations typically occupied by *R. cascadae*	Total number of pixels on path multiplied by percentage of path covered. Areas with elevation within 2 standard deviations of the mean elevation for the sampled populations.	DEM
Moisture/temperature			
compound topographic index (CTI)	Areas characterized by high moisture may prevent desiccation of dispersing frogs.	Mean CTI along path. Index of wetness based on slope and upstream catchment size.	DEM
curvature	Concave surfaces with high moisture may prevent desiccation of dispersing frogs.	Mean curvature along path. Convexity/concavity of surface.	DEM
insolation	Gene flow may be higher across areas receiving more solar radiation, which may be relatively warm and free of persistent ice cover.	Mean insolation for active season (June-September) along path.	DEM
Cover			
forest	Heavily forested areas may reduce gene flow.	Total number of pixels on path multiplied by percentage of path covered.	NWGAP
meadow	Gene flow may be less restricted in habitats typically occupied by *R. cascadae.*	Total number of pixels on path multiplied by percentage of path covered.	NWGAP
open	Gene flow may be less restricted in open habitats, such as shrub and grassland.	Total number of pixels on path multiplied by percentage of path covered.	NWGAP
rock-ice	Barren terrain may reduce gene flow.	Total number of pixels on path multiplied by percentage of path covered.	NWGAP
river	Large rivers may reduce gene flow.	Total number of pixels on path multiplied by percentage of path covered.	WR

*Source data: DEM is a 30 m digital elevation model from the National Elevation Dataset (United States Geological Survey). NWGAP is a landcover dataset by the Northwest Gap Analysis Program (United States Geological Survey). WR is the Washington Rivers dataset from the Washington State Department of Ecology.

Table 4.3 Results of the Random Forests analyses for both 'within' and 'between' genetic group datasets. For each dataset, the best fitting model is indicated in bold type. Variables retained in each model are listed in decreasing order of importance (MIR value). Mean square error (MSE) is the sum of squared residuals divided by the sample size.

Dataset	Structural connectivity model	Pseudo-r^2	MSE	Variables retained
Within				
	Linear	0.42	0.00211	*distance, forest, river, meadow, opt-elev, slope, rock-ice, valley, CTI, ridge*
	Pond	0.38	0.00226	*distance, river, sol, opt-elev, meadow*
	Stream	**0.55**	**0.00164**	***forest, distance, meadow, opt-elev, valley, sol, curv***
Between				
	Linear	0.63	0.00201	*distance, rock-ice, valley, forest, ridge, opt-elev, meadow, sol, open, slope*
	Pond	**0.74**	**0.00145**	*Rock-ice, valley, ridge, CTI, slope, distance, river, ERR*
	Stream	0.72	0.00155	*distance, rock-ice, forest, ridge, sol, meadow, river*

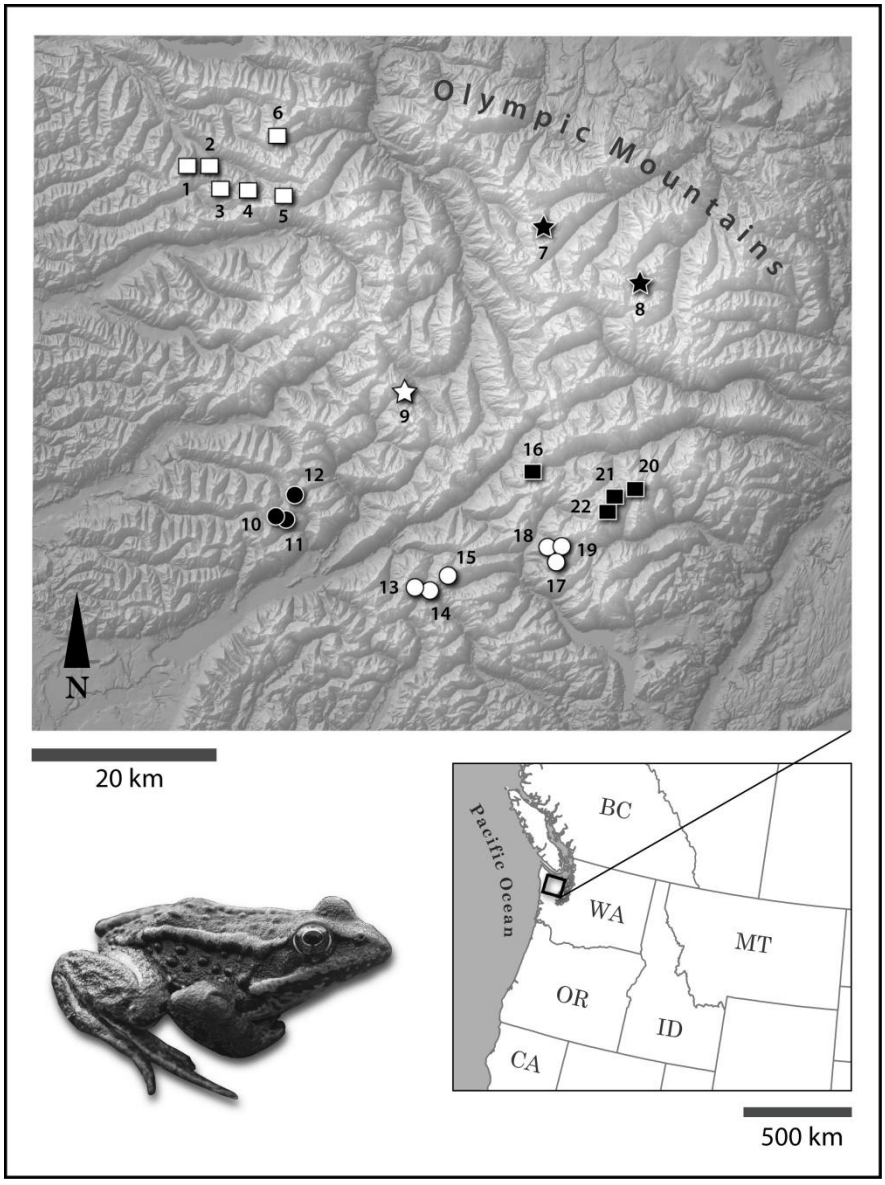

Figure 4.1. Map of *Rana cascadae* population sampling localities in the Olympic Mountains. Populations are numbered as in Table 4.1. Symbols are used to differentiate the six genetic groups of populations identified in the TESS analysis.

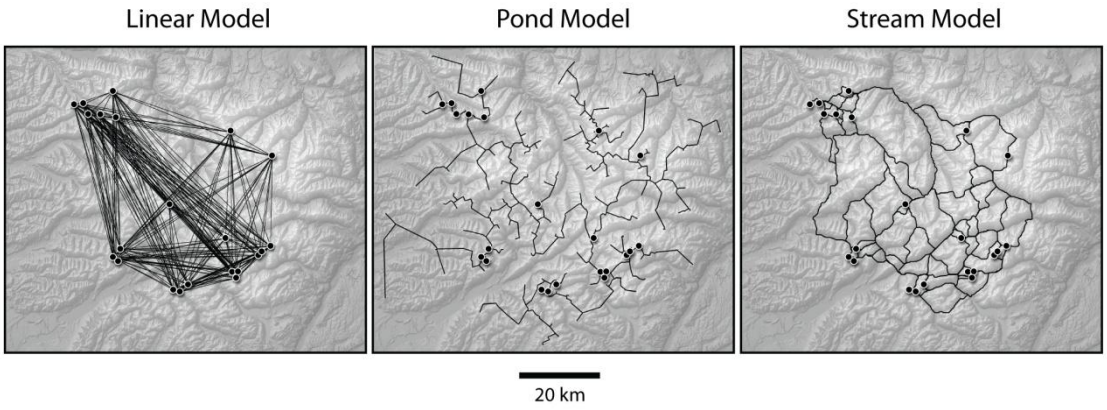

Figure 4.2. Structural connectivity models for 22 *Rana cascadae* populations in the Olympic Mountains. The Linear Model is based on simple Euclidean paths between all pairs of populations. Pairwise paths in the Pond Model follow the branches (i.e. edges) of a minimum spanning tree (MST) that was constructed by connecting all pond habitats above 400 m in elevation. The entire MST is shown in the middle panel. The Stream Model is based on pairwise paths that closely follow stream/river drainages.

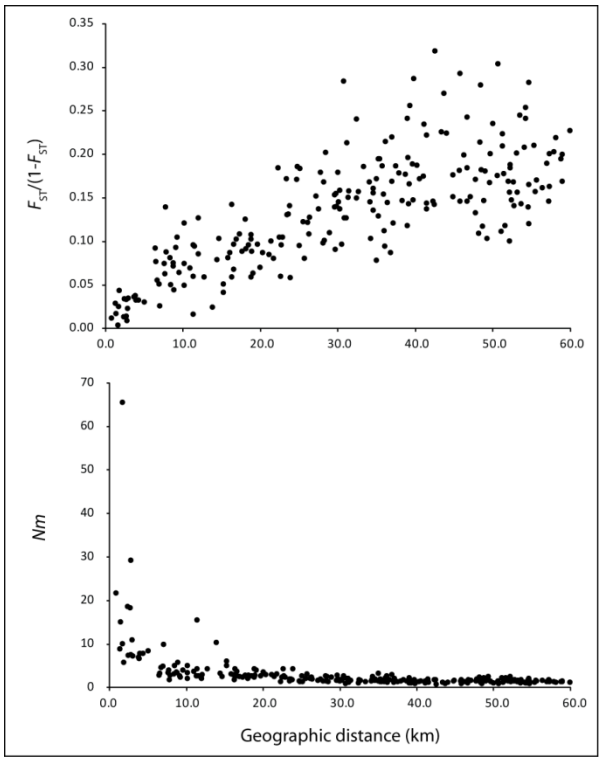

Figure 4.3. Plots of genetic isolation by distance and migration by distance. The top panel is a plot of genetic differentiation versus Euclidean geographic distance. The positive, linear relationship indicates a pattern of isolation by distance. The bottom panel is a plot of migration rate versus geographic distance. The average number of migrants per generation (*Nm*) is very small for populations separated by more than approximately 15 km.

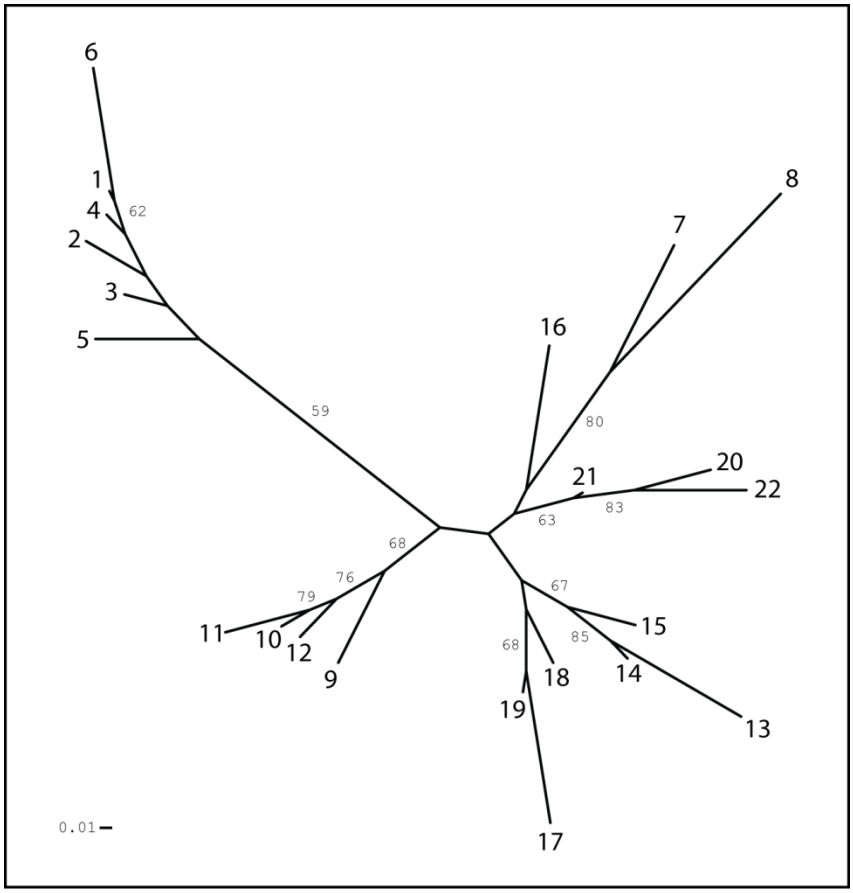

Figure 4.4. Neighbor-joining tree of genetic distances between all 22 *Rana cascade* populations in the Olympic Mountains. Population numbers are at the branch tips and bootstrap values greater than 50 (based on 1000 iterations) are indicated on the branches.

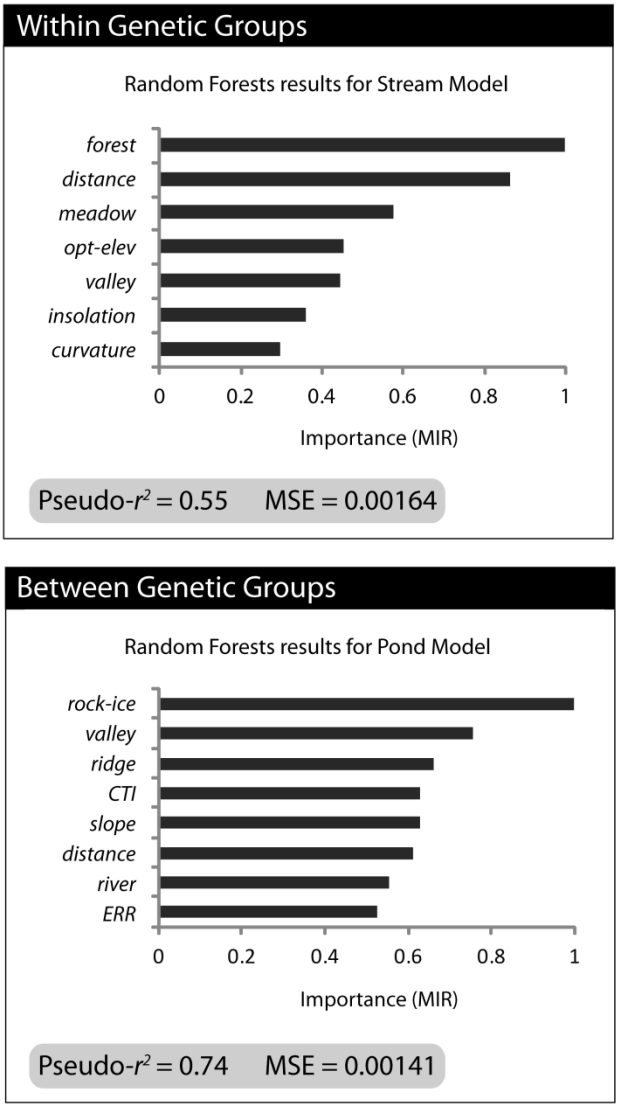

Figure 4.5. Results of the best Random Forests models from analyses for both 'within' and 'between' genetic group datasets. Bar graphs show importance values (model improvement ratios, MIR) for each landscape variable retained in the Random Forests model (maximum possible is 1). Measures of model fit are given as pseudo-r^2 and mean square error (MSE).

CHAPTER 5: CONCLUSION

The evolutionary forces of genetic drift and gene flow play major roles in the ebb and flow of genetic diversity within and among populations. The historical influences of these forces can be estimated and their potential effects can be predicted using the rich theory of population genetics. The gap between the predictions of theory and our understanding of biological reality has narrowed as the analytical methods of population genetics have become more refined and powerful. At the same time, the accessibility and information content of molecular genetic data have continued to increase dramatically, further improving the ability of researchers to address important questions of how evolution proceeds in wild populations.

Nevertheless, the details of how forces such as genetic drift and gene flow influence most species still remain unknown. Differences among species in habitats, physiology, life history, and behavior have likely resulted in a multitude of different patterns of genetic drift (and thus, effective population sizes) and gene flow. To truly understand the importance of these processes within species and how they vary among species, they must be investigated empirically. The challenge is to obtain reliable estimates of N_e and genetic differentiation and then link these to intrinsic (e.g., life history) and extrinsic factors (e.g., landscape barriers to gene flow) that influence genetic drift and gene flow.

In my doctoral research, I have worked towards finding such links for populations of several ranid frog species in North America. In Chapters 2 and 3, I used state of the art methods to obtain estimates of N_e and then used these estimates to

investigate factors that may influence this parameter in pond-breeding ranid frogs. In Chapter 2, I hypothesized that the extrinsic factor of egg mass mortality, a consequence of environmental conditions, would inflate variance in family size for the Oregon spotted frog (*Rana pretiosa*) and thus cause a reduction in the effective number of breeding adults (N_b). This reduction could ultimately lead to a reduction in N_e. The data did not support this hypothesis. Egg mass mortality may not have a strong influence on N_e in *R. pretiosa*. However, more data are certainly needed to make any conclusions regarding this relationship. The research presented in Chapter 2 is among the few studies that have used genetic estimates of effective size at different life stages to investigate the causes of N_e reduction relative to N.

In Chapter 3, I derived directly-comparable N_e estimates for an unprecedented number of populations (90), in order to determine the typical ranges of this parameter in four frog species. I also tested for correlations between N_e in these species and several geographic factors, another way of approaching the question of what factors influences N_e in frog populations. I found intriguing evidence that N_e in two of the species varies predictably across geographic gradients. Although it is unclear whether these relationships are due to elevation differences or to the evolutionary histories of the populations, there are clearly strong patterns that warrant further investigation.

In Chapter 4, I focused on finding links between extrinsic landscape factors and gene flow in the Cascades frog (*Rana cascadae*). I applied used powerful and recently-introduced landscape genetics approach to identify landscape features that are associated with the genetic differentiation of populations. My results improve our

knowledge of how gene flow in this species is influenced by the landscape between populations. I found that using Euclidean distances to model connectivity between populations is an oversimplification; more biologically realistic models incorporating landscape features explained more of the variation in genetic distances between populations. Landscape genetics studies such as this have only recently become feasible, with the availability of high-resolution geographic and genetic data and the computational tools needed to analyze them.

In conclusion, the results of my doctoral research lead to an increased understanding of effective population size in ranid frogs and of the environmental factors that influence population structure in *R. cascadae*. These studies lay a foundation for further research on the specific factors that influence genetic drift and gene flow in these and similar species. For example, I could replicate the method used in Chapter 2 across multiple populations of *R. pretiosa* or one of the other species to make more robust conclusions about what causes N_e to be so small relative to census population size in these frogs. For *R. cascadae*, I could gather genetic and geographic data from an independent set of populations to test to generality of the landscape genetics models I constructed in Chapter 4. I could also combine these models with climate change prediction models to forecast the population structure of *R. cascadae* under alternative climate scenarios. As each of these species is of conservation concern, my aim would be to conduct research that addresses questions of direct relevance to their long-term preservation, in addition to understanding their basic biology.

Bibliography

Araki H, Waples RS, Ardren WR, Cooper B, Blouin MS (2007) Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms. Mol Ecol 16:953-966

Aspi J, Roininen E, Kiiskilä J, Ruokonen M, Kojola I, Bljudnik L, Danilov P, Heikkinen S, Pulliainen E (2008) Genetic structure of the northwestern Russian wolf populations and gene flow between Russia and Finland. Conservat Genet doi:10.1007/s10592-008-9642-x.

Beebee T (2005) Conservation genetics of amphibians. Heredity 95:423-427

Beebee TJC (2009) A comparison of single-sample effective size estimators using empirical toad (*Bufo calamita*) population data: genetic compensation and population size-genetic diversity correlations. *Molecular Ecology* **18**, 4790-4797.

Begon M (1979) Investigating animal abundance: capture-recapture for biologists University Park Press, Baltimore, MD.

Berthier P, Beaumont MA, Cornuet J-M, Luikart G (2002) Likelihood-Based Estimation of the Effective Population Size Using Temporal Changes in Allele Frequencies: A Genealogical Approach. Genetics 160:741-751

Berven KA, Grudzien TA (1990) Dispersal in the wood frog (*Rana sylvatica*): Implications for genetic population structure. Evolution 44, 8:2047-2056

Berven KA (1995) Population regulation in the wood frog, *Rana sylvatica*, from three diverse geographic localities. Aust J Ecol. 20:385–392

Blouin MS (2002) Genetic data for recovery of the Oregon Spotted Frog, *Rana pretiosa*, in the Western United States, Final report for USGS FY 2000 Species at Risk Program project. USGS.

Blouin MS (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends Ecol Evol 18:503-511

Brede EG, Beebee TJ (2006) Large variations in the ratio of effective breeding and census population sizes between two species of pond-breeding anurans. Biol J Linn Soc 89:365-372

Breiman L (2001) Random Forests. *Machine Learning* **45**, 5-32.

Briggs JL (1987) Breeding Biology of the Cascade Frog, *Rana cascadae*, with Comparisons to *R. aurora* and *R. pretiosa*. Copeia 1987:241-245

Brown C (1997) *Habitat structure and occupancy patterns of the montane frog, Rana cascadae, in the Cascades Range, Oregon, at multiple scales: Implications for population dynamics in patchy landscapes*, Oregon State UniversityEditor.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, et al. (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**, 455-457.

Chan LM, Zamudio KR(2009) Population differentiation of temperate amphibians in unpredictable environments. *Molecular Ecology* **18**, 3185-3200.

Charlesworth B (2009) Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* **10**, 195-205.

Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes* **7**, 747-756.

Clark RW, Brown WS, Stechert R, Zamudio KR (2008) Integrating individual behaviour and landscape genetics: the population structure of timber rattlesnake hibernacula. *Molecular Ecology* **17**, 719-730.

Collins JT, Taggart TW (2009) Standard common and current scientific names for North American amphibians, turtles, reptiles, and crocodilians. Sixth Edition. The Center for North American Herpetology.

Corkran C, Thoms C (2006) *Amphibians of Oregon, Washington, and British Columbia* Lone Pine Publishing, Auburn, WA.

Crouch WB, Paton PWC (2000) Using egg-mass counts to monitor wood frog populations. Wildlife Wildl Soc Bull 28:895-901

Crow JF, Morton NE (1955) Measurement of gene frequency drift in small populations. Evolution 9:202-214

Cushman KA, Pearl CA (2007) A conservation assessment for the Oregon spotted frog (*Rana pretiosa*). USDA Forest Service and USDI Bureau of Land Management, Oregon and Washington.

Cutler DR, Edwards TC, Beard KH, et al. (2007) Random Forests for classification in ecology. *Ecology* **88**, 2783-2792.

Davis AB, Verrell PA (2005) Demography and reproductive ecology of the Columbia spotted frog (*Rana luteiventris*) across the Palouse. *Canadian Journal of Zoology* **83**, 702-711.

De'ath G, Fabricius KE (2000) Classification and regression tress: a powerful yet simple technique for ecological data analysis. *Ecology* **81**, 3178-3192.

Dieringer D, Schlotterer C (2003) microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* **3**, 167-169.

Durrant C, Beebee T, Greenaway F, Hill D (2008) Evidence of recent population bottlenecks and inbreeding in British populations of Bechstein's bat, *Myotis bechsteinii*. Conservat Genet doi:10.1007/s10592-008-9639-5. http://dx.doi.org/10.1007/s10592-008-9639-5

Easteal S (1985) The ecological genetics of introduced populations of the giant toad *Bufo marinus*. II. Effective population size. *Genetics* **110**, 107-22.

England P, Cornuet J, Berthier P, Tallmon D, Luikart G (2006) Estimating effective population size from linkage disequilibrium: severe bias in small samples. Conservat Genet 7, 303-308.

Evanno G, Regnauts S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Federal Register (2009) Endangered and Threatened Wildlife and Plants; 90-Day Finding on a Petition to List the Northern Leopard Frog (*Lithobates* [=Rana] *pipiens*) in the Western United States as Threatened. *Federal Register* **74** (207), 55525-26.

Fagre DB, Peterson DL, Hessl AE (2003) Taking the Pulse of Mountains: Ecosystem Responses to Climatic Variability. *Climatic Change* **59**, 263-282.

Fellers GM, Pope K, stead JE, Koo M, Welsh H (2008) Turning population trend monitoring into active conservation: Can we save the Cascades frog (*Rana cascadae*) in the Lassen region of California? *Herpetological Conservation and Biology* **3**, 28-39.

Ficetola GF, Padoa-Schioppa E, Wang J, Garner TWJ (2009) Polygyny, census and effective population size in the threatened frog, *Rana latastei*. *Animal Conservation* **9999**.

Frankham R (1995) Effective population size/adult population size ratios in wildlife: A review. Genet Res 66:95-107

Frankham R (1996) Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology* **10**, 1500-1508.

Fraser DJ, Hansen MM, Østergaard S, *et al.*. (2007) Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Molecular Ecology* **16**, 3866-3889.

Funk WC, Blouin MS, Corn P, *et al.*. (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology* **14**, 483-96.

Funk WC, Cannatella DC, Ryan MJ (2009) Genetic divergence is more tightly related to call variation than landscape features in the Amazonian frogs *Physalaemus petersi* and *P. freibergi*. *Journal of Evolutionary Biology* **22**, 1839-1853.

Funk WC, Pearl CA, Draheim HM, *et al.*. (2008) Range-wide phylogeographic analysis of the spotted frog complex (*Rana luteiventris* and *Rana pretiosa*) in northwestern North America. *Molecular Phylogenetics and Evolution* **49**, 198-210.

Funk WC, Tallmon D, Allendorf F (1999) Small effective population size in the long-toed salamander. *Mol Ecol* **8**, 1633-40.

Garwood J (2009) *Spatial Ecology of the Cascades Frog: Identifying Dispersal, Migration, and Resource Uses at Multiple Spatial Scales*, Humbolt State University.

Giordano AR, Ridenhour BJ, Storfer A (2007) The influence of altitude and topography on genetic structure in the long-toed salamander (*Ambystoma macrodactulym*). *Molecular Ecology* **16**, 1625-1637.

Goudet J (1995) FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *J Hered* **86**, 485-486.

Hayes MP (1997) Status of the Oregon spotted frog (*Rana pretiosa*) in the Deschutes Basin and selected other systems in Oregon and northeastern California with a rangewide synopsis of the species' status. Nature Conservancy under contract

to the U.S. Fish and Wildlife Service with assistance from the Oregon Department of Fish and Wildlife, final report, Portland, Oregon.

Hedrick P (2005) A standardized genetic differentiation measure. *Evolution* **59**, 1633-1638.

Herreid II CF, Kinney S (1966) Survival of Alaskan Woodfrog (*Rana Sylvatica*) Larvae. *Ecology* **47**, 1039-1041.

Hill W (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetical Research*, 209-216.

Hillis DM, Wilcox TP (2005) Phylogeny of the New World true frogs (*Rana*). *Molecular Phylogenetics and Evolution* **34**, 299-314.

Hoffman E, Blouin MS (2004a) Evolutionary history of the northern leopard frog: reconstruction of phylogeny, phylogeography, and historical changes in population demography from mitochondrial DNA. *Evolution* **58**, 145-59.

Hoffman E, Blouin MS (2004b) Historical data refute recent range contraction as cause of low genetic diversity in isolated frog populations. *Mol Ecol* **13**, 271-6.

Hoffman E, Schueler F, Jones A, Blouin MS (2006) An analysis of selection on a colour polymorphism in the northern leopard frog. *Molecular Ecology* **15**, 2627-2641.

Hoffman EA, Ardren WR, Blouin MS (2003) Nine polymorphic microsatellite loci for the northern leopard frog (*Rana pipiens*). *Molecular Ecology Notes* **3**, 115-116.

Hoffman EA, Schueler FW, Blouin MS (2004) Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern leopard frog. *Evolution* **58**, 2536-2545.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806.

Janzen DH (1967) Why Mountain Passes are Higher in the Tropics. *The American Naturalist* **101**, 233.

Jehle R, Arntzen J, Burke T, Krupa A, Hodl W (2001) The annual number of breeding adults and the effective population size of syntopic newts (*Triturus cristatus*, *T. marmoratus*). *Mol Ecol* **10**, 839-50.

Jennings M, Hayes M (1994) *Amphibian and Reptiles Species of Special Concern in California*.

Jones L, Leonard W, Olson D (2005) *Amphibians of the Pacific Northwest* Seattle Audubon Society, Seattle, WA.

Jorde PE, Ryman N (1995) Temporal Allele Frequency Change and Estimation of Effective Size in Populations with Overlapping Generations. Genetics **139**:1077-1090.

Koscinski D, Yates AG, Handford P, Lougheed SC (2009) Effects of landscape and history on diversification of a montane, stream-breeding amphibian. *Journal of Biogeography* **36**, 255-265.

Langella O (1999) Populations 1.2.30: a population genetic software. CNRS UPR9034. Available at http://bioinformatics.org/~tryphon/populations/

Laurila A, Seppa P (1998) Multiple paternity in the common frog (*Rana temporaria*): genetic evidence from tadpole kin groups. Biol J Linn Soc 63:221-232.

Leclerc E, Mailhot Y, Mingelbier M, Bernatchez L (2008) The landscape genetics of yellow perch (*Perca flavescens*) in a large fluvial ecosystem. *Molecular Ecology* **17**, 1702-1717.

Lee-Yaw JA, Davidson A, McRae BH, Green DM (2009) Do landscape processes predict phylogeographic patterns in the wood frog? *Molecular Ecology* **18**, 1863-1874.

Leonard W, Hallock L, McAllister KR (1997) Natural history notes: *Rana pretiosa* (Oregon spotted frog). Behavior and reproduction. Herpetological Review 28:86.

Liaw A, Wiener M (2002) Classification and regression by Random-Forest. *R News* **2**, 18-22.

Licht L (1969) Comparative breeding behavior of the red-legged frog (*Rana aurora aurora*) and the western spotted frog (*Rana pretiosa pretiosa*) in southwestern British Columbia. *Canadian Journal of Zoology* **47**, 1287-1299.

Licht LE (1971) Breeding Habits and Embryonic Thermal Requirements of the Frogs, *Rana Aurora Aurora* and *Rana Pretiosa Pretiosa*, in the Pacific Northwest. Ecology 52:116-124

Licht L (1974) Survival of embryos, tadpoles, and adults of the frogs *Rana aurora aurora* and *Rana pretiosa pretiosa* sympatric in southwestern British Columbia. *Can J Zool* **52**, 613-27.

Luikart G, Ryman N, Tallmon D, Schwartz M, Allendorf F (2010) Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics* **11**, 355-373.

Manel S, Schwartz M, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution* **18**, 189-197.

Merrell DJ (1968) A Comparison of the estimated size and the "effective size" of breeding populations of the leopard frog, *Rana pipiens*. Evolution **22**:274-283

McAllister KR, Leonard WP (1997) Washington State status report for the Oregon Spotted Frog, p. 38. Washington Department of Fish and Wildlife, Olympia.

Monsen K, Blouin MS (2003) Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. *Molecular Ecology* **12**, 3275-86.

Monsen K, Blouin MS (2004) Extreme isolation by distance in a montane frog *Rana cascadae*. *Conservation Genetics* **5**, 827-835.

Mullen LB, Woods HA, Schwartz M, Sepulveda AJ, Lowe W (2010) Scale-dependent genetic structure of the Idaho giant salamander (*Dicamptodon aterrimus*) in stream networks. *Molecular Ecology* **19**, 898-909.

Murphy MA, Evans JS, Storfer A (2010) Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology* **91**, 252-261.

Naiman RJ, Beechie TJ, Benda LE, et al. (1992) Watershed Management: Balancing Sustainability with Environmental Change. In: *Watershed Management: Balancing Sustainability with Environmental Change* (ed. Naiman RJ), pp. 127-188. Springer-Verlag, New York, USA.

Nei M, Tajima F (1981) Genetic drift and estimation of effective populatoin size. *Genetics* **98**, 625-640.

Nussbaum RA, Brodie Jr. ED, Storm RM (1983) *Amphibians and Reptiles of the Pacific Northwest* University of Idaho Press, Moscow, Idaho.

Olson DH, Anderson PD, Frissell CA, Welsh Jr. HH, Bradford DF (2007) Biodiversity management approaches for stream-riparian areas: Perspectives for Pacific Northwest headwater forests, microclimates, and amphibians. *Forest Ecology and Management* **246**, 81-107.

Olson DH, Leonard WP (1997) Amphibian inventory and monitoring: a standardized approach for the PNW. In: *Sampling Amphibians in Lentic Habitats* (eds. Olson DH, Leonard WP, Bury RB), pp. 1-21. Society for Northwestern Vertebrate Biology, Olympia, WA.

Pavlacky Jr. D, Goldizen AW, Prentis PJ, Nicholls JA, Lowe AJ (2009) A landscape genetics approach for quantifying the relative influence of historic and contemporary habitat heterogeneity on the genetic connectivity of a rainforest bird. *Molecular Ecology* **18**, 2945-2960.

Pearl CA, Adams MJ (2005) Amphibian Declines: The Conservation Status of United States' Species. In: *Amphibian Declines: The Conservation Status of United States' Species* (ed. Lannoo M), pp. 538-540. University of California Press, Berkeley, CA.

Phillipsen IC, Bowerman J, Blouin MS (2009) Effective number of breeding adults in Oregon spotted frogs (*Rana pretiosa*): genetic estimates at two life stages. *Conservation Genetics*. **11**, 737-745.

Pike RJ, Wilson SE (1971) Elevation-Relief Ratio, Hypsometric Integral, and Geomorphic Area-Altitude Analysis. *Geological Society of America Bulletin* **82**, 1079-1084.

Prim R (1957) Shortest connection networks and some generalizations. *Bell Systems Technical Journal* **36**, 1389-1401.

Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-59.

Quéméré E, Crouau-Roy B, Rabarivola C, EEL JR, Chikhi L (2010) Landscape genetics of an endangered lemur (*Propithecus tattersalli*) within its entire fragmented range. *Molecular Ecology* **19**, 1606-1621.

R Development Core Team (2009) *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. J Hered 86:248-249

Richards-Zawacki CL (2009) Effects of slope and riparian habitat connectivity on gene flow in an endangered Panamanian frog, *Atelopus varius*. *Diversity and Distributions* **15**, 796-806.

Rorabaugh, JC (2005) *Rana pipiens* Schreber, 1782, Northern leopard frog. In: *Amphibian Declines: The Conservation Status of United States Species* (ed. Lannoo MJ), pp. 570-577. University of California Press, Berkeley.

Sappington JM, Longshore KM, Thompson DB (2007) Quantifying Landscape Ruggedness for Animal Habitat Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert. *Journal of Wildlife Management* **71**, 1419-1426.

Schmeller D, Merila J (2007) Demographic and Genetic Estimates of Effective Population and Breeding Size in the Amphibian *Rana temporaria*. *Conservation Biology* **21**, 142-151.

Schwartz MK, Copeland JP, Anderson NJ, et al. (2009) Wolverine gene flow across a narrow climatic niche. *Ecology* **90**, 3222-3232.

Scribner KT, Arntzen JW, Burke T (1997) Effective number of breeding adults in *Bufo bufo* estimated from age-specific variation at minisatellite loci. Mol Ecol 6:701-712.

Seburn D, Seburn C (2000) Conservation priorities for the amphibians and reptiles of Canada. Prepared for World Wildlife Fund Canada and the Canadian Amphibian and Reptile Conservation Network.

Semlitsch RD (2008) Differentiating Migration and Dispersal Processes for Pond-Breeding Amphibians. *Journal of Wildlife Management* **72**, 260-267.

Soule M (1976) Allozyme variation, its determinants in space and time. In: *Molecular Evolution* (ed. Ayala F), pp. 46-59. Sinaur Associates, Sunderland, MA.

Spear S, Peterson C, Matocq M, Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology* **14**, 2553-64.

Spear SF, Storfer A (2008) Landscape genetic structure of coastal tailed frogs (*Ascaphus truei*) in protected vs. managed forests. *Molecular Ecology* **17**, 4642-4656.

Stebbins R (2003) *A Field Guide to Western Reptiles and Amphibians* Houghton Mifflin, New York.

Stuart SN, Chanson JS, Cox NA, *et al.* (2004) Status and Trends of Amphibian Declines and Extinctions Worldwide. Science 306:1783 - 1786

Sugg DW, Chesser RK (1994) Effective population size with multiple paternity. Genetics 137:1147–1155.

Tallmon D, Koyuk A, Luikart G, Beaumont M (2008) onesamp: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources* **8**, 299-301.

Thompson R, Whitlock K, Bartlein P, Harrison S, Spaulding W (1993) Climatic Changes in the Western United States since 18,000 yr B.P. In: *Global Climates since the Last Glacial Maximum* (eds. Wright H, Kutzbach J, Webb T, *et al.*.). University of Minnesota Press, Minneapolis.

Urban D, Keitt T (2001) Landscape connectivity: a graph-theoretic perspective. *Ecology* **82**, 1205-1218.

U.S. Fish and Wildlife Service (2004) Species assessment and listing priority assignment form, *Rana luteiventris*, Columbia spotted frog (Great Basin DPS).

U.S. Fish and Wildlife Service (2005) Species assessment and listing priority assessment form. Western Washington Fish and Wildlife Office, Lacey, Washington. Available at http://ecos.fws.gov/docs/candforms_pdf/r1/D02A_V01.pdf

Vandergast A, Bohonak A, Weissman D, Fisher R (2007) Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history: phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: *Stenopelmatus*). *Molecular Ecology* **16**, 977-992.

Vergara P, Marquet P (2007) On the seasonal effect of landscape structure on a bird species: the thorn-tailed rayadito in a relict forest in northern Chile. *Landscape Ecology* **22**, 1059-1071.

Vieites D, Nieto-Román S, Barluenga M, *et al.* (2003) Post-mating clutch piracy in an amphibian. Nature Rev. Genet 4:598-612

Vignieri S (2005) Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). *Mol Ecol* **14**, 1925-37.

Vredenburg V, Tunstal T, Bingham R, et al. (2004) Patterns of habitat use and movement of Rana muscosa in the northern Sierra Nevada with comparisons to populations in the southern Sierra Nevada, with additional information on the biogeography of the species. CA Department of Fish and Game.

Waldman B, McKinnon JS (1993) The Natural History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives. In: *The Natural History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives*. University of Chicago Press, Chicago.

Wang IJ (2009) Fine-scale population structure in a desert amphibian: landscape genetics of the black toad (*Bufo exsul*). *Molecular Ecology* **18**, 3847-3856.

Wang IJ, Summers K (2010) Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology* **19**, 447-458.

Wang J (2001) A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genet Res* **78**, 243-57.

Wang J (2005) Estimation of effective population sizes from data on genetic markers. *Philos Trans R Soc Lond B Biol Sci* **360**, 1395-409.

Wang J, Whitlock MC (2003) Estimating effective population size and migration rates From genetic samples over space and time. *Genetics* **163**, 429-446.

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**, 379-91.

Waples RS (1991) Genetic methods for estimating the effective size of cetacean populations. Report of the International Whaling Commission (special issue) 13:279–300

Waples RS (2002) Definition and estimation of effective population size in the conservation of endangered species. In: *Population Viability Analysis* (eds. Beissinger SR McCullough D), pp. 147-168. University of Chicago Press, Chicago.

Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology* **14**, 3335-3352.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics* **7**, 167-184.

Waples RS (2010) Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. *Molecular Ecology Resources* DOI 10.1111/j.1755-0998.2010.02876.x

Waples RS, Do C (2008) ldne: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources* **8**, 753-756.

Waples RS, Do C (2009) Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* **3**, 244-262.

Waples RS, Yokota M (2007) Temporal Estimates of Effective Population Size in Species With Overlapping Generations. *Genetics* **175**, 219-233.

Watson JW, McAllister KR, Pierce DJ, Alvarado A (2000) Ecology of a remnant population of Oregon spotted frogs (*Rana pretiosa*) in Thurston County, Washington. Washington Department of Fish and Game, Olympia, WA.

Weir B, Cockerham C (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**, 1358-1370.

Wells K (1977) The social behaviour of anuran amphibians. Anim Behav 25:666-693

White GC, Burnham KP, Otis DL, Anderson DR (1978) *Users manual for program CAPTURE* Utah State University Press, Logan, UT.

Wilson G, Fulton TL, Kendell K, *et al.*. (2008) Genetic diversity and structure in Canadian northern leopard frog (*Rana pipiens*) populations: implications for reintroduction programs. *Canadian Journal of Zoology* **86**, 863-874.

Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114-138.

Zeisset I, Beebee TJC (2003) Population genetics of a successful invader: the marsh frog *Rana ridibunda* in Britain. *Molecular Ecology* **12**, 639-646.

Zhan A, Li C, Fu J (2009) 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**, 17.

Zhao S, Dai Q, Fu J (2009) Do rivers function as genetic barriers for the plateau wood frog at high elevations? *Journal of Zoology* **279**, 270-276.

Zolbrod A, Peterson DL (1999) Response of high-elevation forests in the Olympic Mountains to climatic change. *Canadian Journal of Forest Research* **29**, 1966-1978.

APPENDICES

APPENDIX A: Supplementary Materials, Chapter 3

Table A.1. Collection locality information for populations of the four frog species used in this study. Sample size after exclusion of monomorphic loci and individuals with missing data is given by *n*. Spatial data are given by latitude (Lat), longitude (Lon), and elevation (Elev). Measures of genetic diversity are given as expected heterozygosity (H_e) and allelic richness (*AR;* rarefied to common sample size of 15). Separate, temporally-spaced samples taken from a single population have names ending with "-A" (first sample) or "-B" (second sample). The collection year is shown as a superscript for each of these samples. H_e and *AR* are not reported for populations that had too much missing data or small sample size.

Species	Population	Orig. name*	*n*	Lat	Lon	Elev (m)	H_e	*AR*
Rana pretiosa								
	RP1-A^{1999}	DC	28	46.9667	-123.0000	42	0.472	3.07
	RP1-B^{2006}	DC	33	46.9667	-123.0000	42	0.446	2.98
	RP2	BC	26	46.8833	-122.9167	77	0.491	3.27
	RP3	TL	35	46.0167	-121.5333	596	0.170	1.91
	RP4	CB	35	45.9500	-121.3167	555	0.503	3.98
	RP5	CP	26	45.1373	-121.5690	962	0.140	1.64
	RP6	HL	31	43.9701	-121.7730	1518	0.239	1.87
	RP7	LC	33	43.8030	-121.8738	1451	0.248	2.02
	RP8	LL	37	43.9108	-121.7572	1445	0.225	1.78
	RP9	WI	32	43.7003	-121.7708	1325	0.242	1.85
	RP10-A^{1999}	SR	28	43.8684	-121.4536	1269	0.292	2.28
	RP10-B^{2006}	SR	32	43.8684	-121.4536	1269	0.227	1.88
	RP11†	-	201	43.8512	-121.4474	1269	0.234	1.85
	RP12	LP	31	43.6827	-121.5161	1282	0.281	2.36
	RP13	DL	28	43.6356	-121.8571	1346	0.323	2.38
	RP14	BM	24	43.3916	-121.9539	1443	0.225	2.29
	RP15	GL	26	43.6326	-122.0432	1466	0.197	1.75
	RP16	JC	27	43.1514	-121.5367	1497	0.374	2.54
	RP17	KE	25	42.9625	-121.5856	1381	0.457	2.92
	RP18	KW	13	42.9464	-121.7485	1379	-	-
	RP19	AR	26	42.9333	-121.4833	1387	0.361	2.61
	RP20	WR	25	42.6233	-121.9714	1263	0.378	3.44
	RP21	BL	39	42.2518	-122.2043	1506	0.305	2.83
Rana luteiventris								
	RL1	KSM	28	48.3178	-115.9796	1581	0.499	3.58
	RL2	KFG	26	48.3347	-115.9737	884	0.586	4.25
	RL3	KLB	19	48.3309	-115.9215	785	0.631	4.99
	RL4	KSF	28	48.3267	-115.9187	824	0.550	4.44
	RL5	KUSF	17	48.3241	-115.9232	812	0.618	4.16
	RL6	SSCL	22	48.2347	-115.9177	1485	0.401	3.29
	RL7	MUB	25	47.8899	-115.8350	833	0.652	5.34
	RL8	MMB	50	47.8910	-115.8300	819	0.665	5.04
	RL9	MUBDG	25	47.8903	-115.7964	769	0.619	4.62
	RL10	MHB	18	47.8876	-115.7938	839	0.651	4.95
	RL11	MBAM	25	47.8745	-115.7551	769	0.642	5.08
	RL12	OSOH	22	46.6620	-114.2366	2251	0.444	2.78
	RL13	SWDPE	24	46.6415	-114.2526	2244	0.471	3.40
	RL14	SWDPW	25	46.6406	-114.2534	2241	0.456	3.45
	RL15	SWPLW	19	46.6326	-114.2359	1982	0.425	2.82

	RL16	SWRR	23	46.5746	-114.0867	999	0.421	2.70
	RL17	SWSFS	25	46.6131	-114.2700	2238	0.398	2.56
	RL18	RURC	20	46.0202	-114.4159	2133	0.389	2.40
	RL19	RRCP	21	46.0746	-114.2105	1250	0.505	3.48
	RL20	LFP	20	46.0127	-114.3797	2256	0.318	2.75
	RL21	LBLP	22	46.0069	-114.3668	2139	0.291	2.49
	RL22	LLRCL	18	46.0229	-114.3420	1995	0.344	2.28
	RL23	SKFP	21	45.1121	-114.5989	2484	0.354	2.85
	RL24	SKB	30	45.1202	-114.5826	2652	0.395	2.97
	RL25	TT	30	45.0896	-114.6116	2548	0.261	2.20
	RL26	TA	24	45.0756	-114.6161	2560	0.231	1.74
Rana cascadae								
	RC1	Many Lakes	24	43.8155	-121.9068	1563	0.538	3.07
	RC2	Berkeley Park	18	46.9131	-121.6872	1971	0.490	4.01
	RC3	Gold Lake	31	43.6332	-122.0464	1467	0.736	5.90
	RC4	Todd Lake	23	44.0250	-121.6821	1895	0.553	4.02
	RC5	Benson Lake	23	44.2322	-121.9157	1684	0.513	4.46
	RC6	Reflection Lakes	23	46.7680	-121.7264	1520	0.511	4.20
	RC7	Paradise River	18	46.7786	-121.7368	1498	0.501	4.02
	RC8	North Waldo	23	43.7623	-122.0131	1600	0.656	4.93
	RC9	Melakwa Lake	15	44.1973	-121.9089	1497	-	-
	RC10	Seven Mile	13	42.7161	-122.1278	1448	-	-
	RC11	Elysian Fields	29	46.9435	-121.7554	1792	0.433	3.56
	RC12	McKenzie Pass	28	44.2448	-121.8414	1585	0.535	4.30
	RC13	Waldo Lake	27	43.7623	-122.0131	1664	0.676	4.85
	RC14	Illabot Creek	20	48.4402	-121.3876	1331	0.326	2.17
	RC15	Mt. Ranier	29	46.9160	-121.6531	1922	0.492	4.33
	RC16	Olympic	29	47.9163	-123.7814	1083	0.512	4.19
	RC17	Big Frank	18	42.4422	-122.2416	1780	-	-
	RC18	Crystal Springs	22	43.3123	-122.1404	1275	-	-
	RC19	Breitenbush	27	44.7716	-121.9495	732	0.658	4.97
	RC20	Colby Creek	30	40.1113	-121.4846	1496	0.576	4.52
Lithobates pipiens								
	LP1-A[1971]	NONQ	40	42.9893	-76.7715	117	0.885	12.66
	LP1-B[2001]	NONQ	54	42.9893	-76.7715	117	0.872	12.49
	LP2-A[1971]	MONTZ	41	43.4679	-76.0100	182	0.903	12.46
	LP2-B[2001]	MONTZ	39	43.4679	-76.0100	182	0.907	13.10
	LP3-A[1979]	HAPVY	38	45.0687	-75.6530	86	0.884	12.28
	LP3-B[2001]	HAPVY	44	45.0687	-75.6530	86	0.856	10.67
	LP4-A[1971]	FAIRM	39	43.5180	-79.9970	312	0.874	11.84
	LP4-B[2001]	FAIRM	43	43.5180	-79.9970	312	0.876	11.96
	LP5	CAMPB	41	44.0370	-78.9790	317	0.872	12.34
	LP6	1	36	46.9208	-92.1555	416	0.728	6.12
	LP7	2	36	44.2153	-93.5931	315	0.856	10.64
	LP8	3	30	45.1917	-93.6917	289	0.836	10.09
	LP9	4	37	45.1264	-94.0306	297	0.850	10.24
	LP10	5	32	45.7097	-94.8931	381	0.841	9.82
	LP11	6	26	42.0006	-97.0064	426	0.833	9.16
	LP12	7	33	42.1734	-98.1558	541	0.807	8.32
	LP13	8	23	42.2371	-100.0883	820	0.737	6.66
	LP14	9	25	41.9467	-102.4334	1183	0.582	3.09
	LP15	10	22	41.9652	-102.3390	1181	0.706	5.79
	LP16	11	30	41.9393	-102.3872	1180	0.719	5.95

LP17	12	25	42.4005	-100.7893	919	0.791	8.37
LP18	13	22	42.1385	-111.2623	1877	0.659	6.37
LP19	14	23	42.2568	-112.0131	1451	0.627	5.69
LP20	15	30	42.6198	-113.2837	1280	0.631	5.12
LP21	16	18	42.6120	-113.2473	1279	0.540	4.75
LP22	17	26	49.0500	-116.5017	626	0.422	2.56
LP23	-	29	47.07613	-119.35362	318	0.244	2.08

* Original population names from Blouin et al (*in review*; *R. pretiosa*), Funk et al (2005; *R. luteiventris*), Monsen and Blouin (2003; *R. cascadae*), Hoffman et al (2004; *L. pipiens* populations LP1-5), and Hoffman et al (2006; *L. pipiens* populations LP6-LP22).

† Data for this population was collected in the study of Phillipsen et al (2010).

Table A.2. Microsatellite loci used for each frog species in this study.

Locus	*R. pretiosa*	*R. luteiventris*	*R. cascadae*	*L. pipiens*
RC287			•	
RP3	•	•		
RP15		•		
RP17	•	•	•	
RP22	•			
RP23	•	•		
RP26	•			
RP193	•		•	•
RP385				
RP415	•			•
RP461	•			
SFC104	•			
SFC134	•	•	•	
SFC120	•		•	
SFC128			•	
SFC139		•		
Rpi100				•
Rpi101				•
Rpi103				•
Rpi104				•
Rpi106				•
Rpi107				•
Rp108				•
n =	11	6	6	9

Table A.3. Estimates of effective population size for populations of the four frog species. Estimates were obtained using the single-sample linkage disequilibrium (LD) method and approximate Bayesian computation (ABC) method. For estimates from the latter, the upper Bayesian prior for $\widehat{N}_e$ is given in the subscript. Confidence limits are shown in parentheses. See Table 3.1 for notation used.

Species	Population	$\widehat{N}_{e(LD)}$	$\widehat{N}_{e(ABC,200)}$	$\widehat{N}_{e(ABC,2000)}$
Rana pretiosa				
	RP1-A	44 (19-791)	20 (17-26)	30 (21-79)
	RP1-B	32 (16-95)	20 (17-26)	28 (22-56)
	RP2	47 (19-∞)	23 (19-34)	28 (19-71)
	RP3	41 (7-∞)	20 (13-35)	21 (12-61)
	RP4	-670 (78-∞)	35 (27-55)	45 (29-163)
	RP5	43 (3-∞)	8 (7-13)	5 (4-8)
	RP6	38 (7-∞)	17 (12-25)	30 (18-102)
	RP7	19 (9-49)	12 (10-17)	9 (7-29)
	RP8	30 (4-∞)	32 (19-63)	35 (19-73)
	RP9	21 (6-224)	21 (16-36)	30 (20-93)
	RP10-A	-63 (38-∞)	20 (16-31)	42 (24-136)
	RP10-B	-83 (15-∞)	27 (19-51)	27 (15-66)
	RP11	33 (15-69)	30 (20-55)	34 (19-66)
	RP12	-171 (31-∞)	16 (12-24)	48 (27-153)
	RP13	20 (6-∞)	29 (22-53)	23 (15-57)
	RP14	638 (9-∞)	18 (14-29)	33 (19-128)
	RP15	-496 (8-∞)	14 (10-23)	23 (15-68)
	RP16	16 (8-40)	20 (16-29)	33 (23-93)
	RP17	42 (15-∞)	13 (10-18)	32 (22-104)
	RP18	468 (12-∞)	19 (15-30)	19 (13-45)
	RP19	45 (15-∞)	18 (15-25)	32 (22-97)
	RP20	67 (18-∞)	22 (18-34)	33 (22-95)
	RP21	151 (35-∞)	17 (11-33)	17 (10-40)
Rana luteiventris				
	RL1	27 (6-∞)	16 (11-24)	20 (12-54)
	RL2	-407 (16-∞)	29 (22-46)	27 (18-64)
	RL3	220 (20-∞)	24 (18-40)	32 (21-76)
	RL4	55 (14-∞)	38 (26-82)	35 (19-98)
	RL5	67 (9-∞)	21 (15-37)	32 (19-108)
	RL6	-75 (27-∞)	23 (16-41)	32 (19-102)
	RL7	11 (3-64)	22 (17-35)	24 (14-53)
	RL8	29 (13-100)	42 (27-86)	65 (35-186)
	RL9	20 (12-42)	32 (24-62)	45 (25-153)
	RL10	-70 (21-∞)	20 (15-34)	38 (21-107)
	RL11	80 (13-∞)	33 (25-55)	30 (19-78)
	RL12	3 (2-12)	12 (8-19)	12 (8-27)
	RL13	22 (3-∞)	19 (13-33)	23 (13-69)
	RL14	-58 (21-∞)	24 (17-42)	19 (13-39)
	RL15	101 (1-∞)	12 (9-20)	12 (9-29)
	RL16	4 (1-∞)	8 (6-13)	9 (6-18)
	RL17	21 (4-∞)	15 (10-27)	14 (8-29)
	RL18	4 (1-∞)	14 (10-22)	13 (9-31)
	RL19	17 (3-∞)	24 (17-42)	23 (14-56)
	RL20	91 (8-∞)	24 (17-49)	30 (15-82)

	RL21	2 (1-8)	19 (12-39)	13 (7-31)
	RL22	-42 (3-∞)	21 (13-37)	28 (16-60)
	RL23	10 (1-∞)	22 (15-40)	26 (16-69)
	RL24	9 (3-27)	21 (15-43)	29 (14-74)
	RL25	72 (1-∞)	11 (8-19)	14 (9-42)
	RL26	3 (1-∞)	10 (7-19)	9 (6-27)
Rana cascadae				
	RC1	52 (3-∞)	25 (18-53)	23 (13-59)
	RC2	28 (6-∞)	15 (11-25)	23 (15-55)
	RC3	16 (11-24)	34 (26-57)	54 (35-159)
	RC4	18 (8-72)	28 (21-49)	23 (14-64)
	RC5	-69 (37-∞)	29 (22-53)	37 (23-108)
	RC6	116 (17-∞)	28 (20-51)	51 (28-177)
	RC7	32 (6-∞)	24 (19-39)	30 (19-91)
	RC8	46 (9-∞)	30 (23-56)	40 (27-137)
	RC9	-39 (14-∞)	25 (18-42)	24 (14-71)
	RC10	3 (2-8)	14 (11-26)	18 (12-42)
	RC11	-138 (21-∞)	36 (25-72)	43 (22-151)
	RC12	17 (11-28)	35 (25-64)	36 (21-98)
	RC13	-64 (11761-∞)	33 (25-58)	38 (25-86)
	RC14	-15 (3-∞)	21 (15-41)	27 (15-81)
	RC15	41 (14-∞)	38 (29-64)	47 (28-174)
	RC16	1188 (34-∞)	36 (27-62)	27 (18-64)
	RC17	34 (5-∞)	21 (15-39)	19 (13-44)
	RC18	11 (4-34)	18 (15-29)	17 (12-44)
	RC19	-44 (-103-∞)	28 (20-48)	29 (16-68)
	RC20	7 (4-11)	27 (18-52)	37 (21-128)
Lithobates pipiens				
			$\widehat{N}_{e(ABC,2000)}$	$\widehat{N}_{e(ABC,5000)}$
	LP1-A	-786 (741-∞)	12681 (3728-142875)	68572 (6777-18621956)
	LP1-B	325 (151-∞)	1870 (641-9572)	3093 (860-25497)
	LP2-A	-5446 (245-∞)	209 (111-614)	1769 (602-15274)
	LP2-B	9337 (343-∞)	7737 (2154-64223)	10996 (2253-275552)
	LP3-A	1293 (319-∞)	751 (252-4424)	5525 (1516-38929)
	LP3-B	136 (46-∞)	97 (52-304)	188 (91-684)
	LP4-A	373 (141-∞)	732 (309-2582)	574 (220-2741)
	LP4-B	1087 (241-∞)	2032 (797-13683)	1745 (533-14550)
	LP5	581 (200-∞)	764 (393-2262)	7251 (1720-89068)
	LP6	136 (57-∞)	64 (41-169)	67 (39-197)
	LP7	210 (114-980)	764 (393-2262)	1541 (569-8798)
	LP8	131 (69-744)	572 (263-2401)	789 (307-4134)
	LP9	255 (118-∞)	350 (178-922)	1004 (402-5058)
	LP10	145 (80-543)	183 (100-422)	245 (130-958)
	LP11	55 (34-118)	115 (68-347)	171 (87-622)
	LP12	49 (35-78)	116 (70-324)	143 (71-433)
	LP13	11 (9-15)	33 (22-70)	49 (28-161)
	LP14	220 (3-∞)	16 (12-30)	17 (12-35)
	LP15	168 (33-∞)	48 (30-120)	47 (27-132)
	LP16	36 (22-74)	48 (32-121)	59 (33-156)
	LP17	69 (41-176)	80 (50-189)	120 (64-350)
	LP18	71 (27-∞)	58 (34-169)	77 (40-262)
	LP19	225 (39-∞)	41 (25-99)	44 (26-129)
	LP20	27 (18-48)	41 (27-119)	30 (18-83)

LP21	5 (2-14)	21 (15-46)	30 (20-80)
LP22	9 (3-32)	23 (16-56)	37 (21-135)
LP23	6 (2-25)	25 (15-79)	32 (17-108)

APPENDIX B: Supplementary Materials, Chapter 4

Table B.1. Microsatellite loci information for *Rana cascadae*. Loci are grouped by the multiplex PCR sets.

Locus	Multiplex set	Primer name	Sequence (5'-3')	Volume (µL) of primer in 500 µL primer mix
RP193	A	MB224	CCATTTTCTCTCTGATGTGTGT	1.5
		MB225	TGAAGCAGATCACTGGCAAAGC	1.5
SFC128	A	MB270	AGAAAAGCGGACTTCTGAAAT	5.0
		MB271	AGCCATAATCCCTGTTAAACC	5.0
SFC134	A	MB276	TGGGAAAAGACTCTGTGGT	5.0
		MB277	AGGAAATGTGTGGAAGCAT	5.0
D114	B	D114_F	CCTGGTGCCATTATTTTTTTAG	7.5
		D114_R	TTATCCCGGAGGAGTACAGTC	7.5
D119	B	D119_F	ATGCAGTTTACAGTTTCACACG	2.5
		D119_R	ATCCCCACACACGCTCTA	2.5
D129	B	D129_F	CCAAAGACAGAGGCACTTAG	2.5
		D129_R	TGCTCAGGACCTGTAGGTAG	2.5
D209	B	D209_F	GCACAGGGACACACACATC	2.5
		D209_R	GCTCGGAGATAGGTAGGGG	2.5
D131	C	D131_F	CCTTTGGAGGACGATACAGG	2.5
		D131_R	GCAGACAGTAGCACAGCACAC	2.5
D208	C	D208_F	AGTCCTTCTCCACTTTTTTCTC	2.5
		D208_R	CAGCCTGTTCTGGGTTATT	2.5
RP415	C	MB234	AAGTTTCATTAAAGCAGATT	2.5
		MB235	GGTATATCTTAGGGTTACCT	2.5

STRUCTURE analysis

Methods

As a complement to the TESS analysis of population structure in *Rana cascadae*, we used STRUCTURE (Pritchard *et al.*, 2000) to identify the major genetic groups represented by the 22 sampled populations. For each value of *K* (the hypothesized number of distinct genetic groups) from 1-20, we carried out 20 independent runs under the correlated allele frequencies model allowing admixture. Runs were performed assuming the correlated allele frequencies model and admixture. Each run had a total of 1 million iterations with a burn-in of 50,000 iterations. For each value of *K*, we calculated the mean and standard deviation of $\ln \Pr(X|K)$ (the estimated likelihood of K) across the 20 runs. We applied the ΔK method of Evanno *et al*. (2005) to identify the most likely number of genetic groups. We first performed a STRUCTURE analysis using the entire dataset of 544 individuals. This allowed us to identify the highest level of hierarchical population structure. We then carried out separate analyses on each major group identified by the first analysis (for values of *K* from 1 to 10). Through this iterative procedure we characterized the overall hierarchical structure of the populations.

Results

We identified several levels of hierarchical structure in the 22 *R. cascadae* populations sampled in the Olympic Mountains (see Figure B.1). At the highest level of structure, the northwestern populations form a distinct group and the remaining populations form a second group. Subsequent analysis of these groups identified smaller hierarchical population groups that were geographically clustered. These results are in agreement with those of the TESS analysis.

Table B.2. Pairwise measures of genetic differentiation between *Rana cascadae* populations in the Olympic Mountains. Populations are numbered as in Table 4.1 and Figure 4.1. F_{ST} is given in the lower half of the matrix and D_{ps} is given in the upper half. All pairs of populations were significantly differentiated after Bonferroni correction in Fisher's exact tests *except* for those marked with † (only shown for F_{ST}).

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
1	-	0.25	0.27	0.25	0.35	0.29	0.66	0.64	0.55	0.58	0.58	0.59	0.63	0.59	0.56	0.63	0.58	0.60	0.60	0.64	0.58	0.60
2	0.01†	-	0.28	0.29	0.40	0.37	0.60	0.62	0.55	0.54	0.54	0.57	0.55	0.52	0.50	0.61	0.53	0.56	0.54	0.63	0.53	0.59
3	0.03	0.02†	-	0.28	0.34	0.36	0.59	0.59	0.52	0.54	0.53	0.54	0.54	0.51	0.48	0.58	0.49	0.55	0.53	0.55	0.50	0.57
4	0.02	0.03†	0.03	-	0.30	0.30	0.63	0.60	0.57	0.58	0.58	0.58	0.61	0.56	0.54	0.62	0.52	0.57	0.55	0.59	0.56	0.58
5	0.06	0.07	0.05	0.03	-	0.36	0.58	0.62	0.58	0.55	0.56	0.54	0.61	0.56	0.54	0.58	0.52	0.56	0.56	0.57	0.55	0.54
6	0.05	0.08	0.08	0.05†	0.07	-	0.67	0.65	0.59	0.61	0.61	0.62	0.64	0.58	0.63	0.68	0.59	0.66	0.61	0.67	0.62	0.64
7	0.20	0.18	0.16	0.19	0.17	0.22	-	0.42	0.49	0.50	0.51	0.54	0.58	0.54	0.54	0.48	0.61	0.51	0.52	0.55	0.44	0.53
8	0.23	0.22	0.20	0.21	0.22	0.24	0.11	-	0.53	0.54	0.55	0.61	0.56	0.55	0.54	0.50	0.48	0.49	0.48	0.54	0.46	0.54
9	0.14	0.14	0.12	0.15	0.15	0.18	0.12	0.14	-	0.38	0.37	0.37	0.47	0.41	0.40	0.46	0.53	0.45	0.49	0.51	0.40	0.50
10	0.16	0.15	0.13	0.16	0.13	0.19	0.12	0.16	0.06	-	0.25	0.27	0.52	0.47	0.44	0.42	0.51	0.46	0.49	0.49	0.45	0.48
11	0.16	0.14	0.13	0.16	0.15	0.18	0.13	0.17	0.06	0.01†	-	0.30	0.48	0.46	0.44	0.49	0.53	0.46	0.51	0.52	0.44	0.49
12	0.16	0.14	0.13	0.16	0.13	0.19	0.13	0.18	0.06	0.01†	0.03†	-	0.51	0.47	0.43	0.47	0.52	0.48	0.49	0.48	0.43	0.44
13	0.14	0.10	0.10	0.13	0.15	0.17	0.13	0.15	0.08	0.09	0.08	0.09	-	0.27	0.38	0.51	0.44†	0.39	0.42	0.46	0.38	0.48
14	0.14	0.09	0.09	0.12	0.13	0.16	0.12	0.15	0.07	0.09	0.08	0.08	0.00†	-	0.25	0.46	0.45	0.33	0.36	0.43	0.36	0.47
15	0.14	0.11	0.11	0.13	0.13	0.18	0.12	0.15	0.08	0.09	0.09	0.08	0.04	0.01	-	0.43	0.43	0.28	0.36	0.46	0.39	0.45
16	0.19	0.18	0.15	0.18	0.16	0.23	0.11	0.15	0.12	0.09	0.12	0.11	0.11	0.10	0.09	-	0.51	0.35	0.40	0.39	0.36	0.43
17	0.16	0.14	0.12	0.13	0.14	0.17	0.18	0.14	0.15	0.11	0.13	0.13	0.07	0.07	0.08	0.11	-	0.36	0.27	0.43	0.39	0.47
18	0.16	0.14	0.12	0.15	0.15	0.20	0.12	0.14	0.09	0.08	0.10	0.09	0.04	0.02	0.02	0.05	0.04†	-	0.27	0.36	0.34	0.41
19	0.17	0.14	0.12	0.15	0.15	0.19	0.14	0.13	0.12	0.11	0.12	0.12	0.06	0.05	0.06	0.07	0.02†	0.02†	-	0.39	0.35	0.42
20	0.19	0.18	0.15	0.17	0.17	0.22	0.15	0.16	0.13	0.11	0.13	0.11	0.10	0.09	0.09	0.09	0.09	0.06	0.07	-	0.28	0.30
21	0.14	0.13	0.11	0.14	0.13	0.17	0.09	0.12	0.07	0.08	0.09	0.07	0.06	0.06	0.07	0.06	0.08	0.04	0.06	0.03	-	0.27
22	0.17	0.17	0.14	0.16	0.15	0.20	0.13	0.16	0.11	0.10	0.11	0.09	0.09	0.09	0.10	0.09	0.12	0.07	0.08	0.03	0.03	-

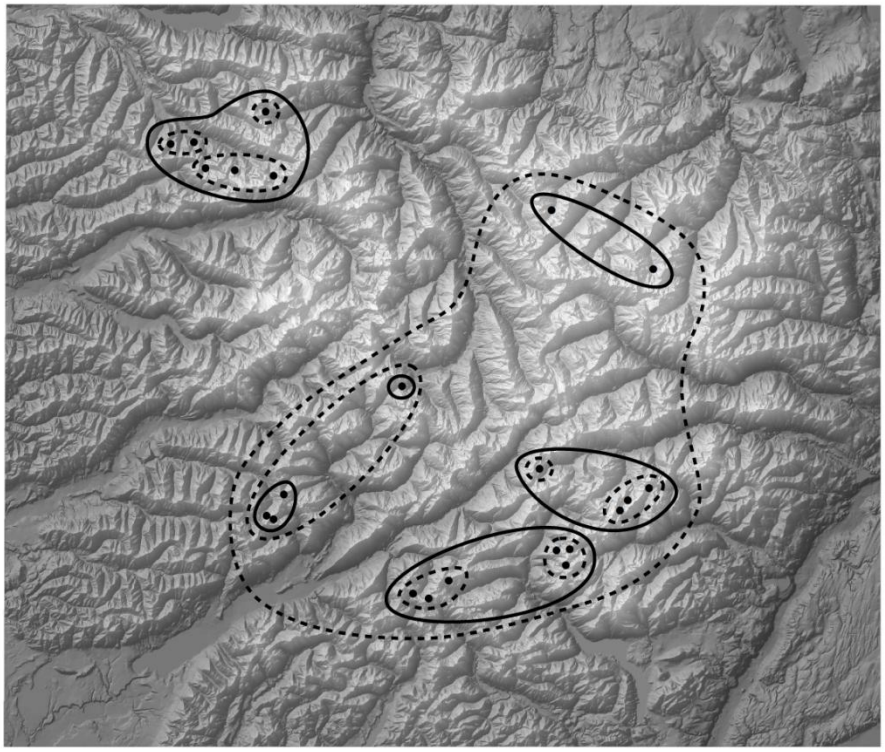

Figure B.1. Hierarchical population groups identified in the TESS and STRUCTURE analysis. The nesting pattern of the groups indicates the hierarchy. For example, the six populations in the northwest part of the Olympic Mountains form a single group at the highest hierarchical level. Within this group, the populations are split into three smaller groups. Groups encircled by thick black lines are the groups identified in the TESS analysis. These are the groups used to define the 'within' and 'between' levels of statistical analysis. Groups encircled by dashed lines were identified with STRUCTURE. There were no population groups that were identified by STRUCTURE but not TESS and vice versa. For population information see Figure 4.1 and Table 4.1.