UNDERSTANDING PARAPATRIC RANGE LIMITS IN THE LONG-TOED SALAMANDER, *AMBYSTOMA MACRODACTYLUM*

by

JULIE ANNE LEE-YAW

B.Sc., Queen's University, 2003

M.Sc., McGill University, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(ZOOLOGY)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2013

ABSTRACT

Understanding geographic range limits is an outstanding challenge in evolutionary ecology. My goal was to characterize and evaluate factors contributing to parapatric borders in the long-toed salamander (*Ambystoma macrodactylum*). Using amplified fragment length polymorphism and mitochondrial data, I tested whether currently-recognized subspecies of long-toed salamander are distinct evolutionary units. My results demonstrate that the long-toed salamander consists of at least four divergent lineages. Discordance between these lineages and current subspecies designations, as well as evidence for a cryptic lineage, emphasize the need to reevaluate existing taxonomy prior to conducting studies of species' range limits.

To further understand the distribution of diversity in this system, I explored the role of climate in shaping lineage boundaries. Using spatial data and ecological niche modeling, I asked whether the boundaries between lineages reflect the limits of their respective climatic niches. My results suggest that the different long-toed salamander lineages are ecologically similar and that suitable climatic space for each lineage exists well-beyond shared borders. Although some contact zones coincide with areas where the average climatic suitability for both lineages is low, sites that are highly suitable for each lineage can be found within these regions in all cases. Thus climatic barriers alone are unlikely driving range limits in this system.

I next examined the role of hybridization in shaping range limits. I characterized fine-scale patterns of genetic structure in a contact zone between two long-toed salamander lineages. To determine whether there is evidence of hybrid dysfunction, I assayed adult feeding performance in the laboratory. I observed reduced feeding performance in populations coinciding with the extent of mitochondrial introgression but not in populations that are more admixed. These results may be relevant for understanding the limits of introgression for some genes, but not all. Thus the study of range limits in the context of hybrid zones may require consideration of factors governing differential rates of introgression across the genome. This dissertation demonstrates the use of multiple lines of investigation to narrow down the most relevant hypotheses for parapatric range limits and highlights the potential for several factors to ultimately be shaping species' range limits.

PREFACE

A version of Chapter 2 has been published as: Lee-Yaw JA and Irwin DE (2012) Large geographic range size reflects a patchword of divergent lineages in the long-toed salamander (*Ambystoma macrodactylum*). *Journal of Evolutionary Biology*, **25**, 2276-2287. I conceived of this project in collaboration with Darren Irwin. I conducted both the fieldwork and genetic data collection for this project, performed all analyses and wrote the original manuscript. Darren provided valuable advice on analyses and contributed important revisions toward the production of the final manuscript.

Chapter 4 was completed in collaboration with Chris Jacobs (M.Sc., Leiden University) and Darren Irwin. I conceived of the experiment and designed the sampling scheme. Chris and I collected the salamanders and decided on the final protocols together with advice from Darren. Chris collected all of the feeding data and cared for the animals in the laboratory. I collected all of the genetic data for this study, conducted the statistical analyses and wrote the manuscript.

All protocols involving the use of animals in this dissertation were approved by the UBC animal care committee (Application Numbers: A07-0632 and A11-0382). Permits for the collection of tissue samples and individuals were obtained from the appropriate government agencies and are as follows: BC Ministry of Environment (Scientific Collection Permits:VI07-40215, VI08-50170, VI09-59560, VI11-68724); Alberta Sustainable Resource Development, Fisheries and Wildlife Management (Scientific Collection Licences: 35183, 39420, 47083 and Permits: 36005, 39419, 47084); Parks Canada (Permit: JNP-2009-2252); Washington Department of Fish and Wildlife (Scientific Collection Permit: 07-377); Idaho Department of Fish and Game (Wildlife Collecting Permit: 071106); Montana Department of Fish, Wildlife and Parks (Scientific Collection Permit: 2008-005); Oregon Department of Fish and Wildlife (Scientific Taking Permit: 001-08).

This research was supported by the following grants: National Sciences and Engineering Research Council of Canada Canadian Graduate Scholarship, Pacific Century Graduate Scholarship (University of British Columbia), Alberta Conservation Association Grant in Biodiversity, Sigma Xi Grant-in-Aid-of-Research and the Alfred Bader Fellowship in memory of Jean Royce (Queen's University).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFLP – Amplified fragment length polymorphism

ANOVA – Analysis of variance

AUC – Area under the curve

bp – base pair(s)

CC – Coastal-Cascade long-toed salamander lineage

COH – Central Oregon Highlands long-toed salamander lineage

cyt b – cytochrome b

GIS – Geographic information systems

IDW – Inverse distance weighting

km – kilometer

mtDNA – mitochondrial DNA

MDS – Multidimensional scaling

MEC – Mass conversion efficiency

NC – North-Central long-toed salamander lineage

PC(A) – Principal components (analysis)

PC – Principal components

RM – Rocky-Mountains long-toed salamander lineage

sig – *Ambystoma macrodactylum sigillatum*

ybp – years before present

ACKNOWLEDGEMENTS

The best part has been the people. I thank my supervisor, Darren Irwin, for his encouragement, generosity and help. I especially appreciate his support of my decision to work on the salamanders—a system so far removed from those of his lab group. I learned a lot from Darren and my work has benefited from his insightful comments and suggestions. I also feel fortunate to have had the opportunity to work with the many great students that he has attracted to his lab: Alan Brelsford, Kira Delmore, David Toews, Haley Kenyon, Alison Porter, Christine Grossen, Katie Kuker, Sampath Seneviratne, Miguel Alcaide and Michelle Chen. My conversations with Alan, Dave and Kira were especially helpful.

I am grateful for the support I received from my supervisory committee: Trish Schulte, Dolph Schluter and Mike Whitlock. My committee meetings were always inspiring and though-provoking. I am particularly grateful to Dolph for the extra time he spent discussing statistics with me; to Trish for several critical discussions about hybrid zones and my fourth chapter in particular; and to Mike, for entertaining my many questions about population genetics and statistics. Mike also took me under wing while Darren was away on sabbatical and generously allowed me to use his environmental chambers for my experiments. I would also like to thank Amy Angert for welcoming me to her lab meetings towards the end of my time at UBC. My interactions with her and her students fueled my excitement for range limits anew.

Several people were instrumental to my success in the field and with data collection. Mark Thompson kindly provided samples and readily shared his enthusiasm for and knowledge of the system with me. Andrew Storfer and Paul Verrell provided information about collecting salamanders in the field and about their care in the laboratory. Andrew also gave me a great opportunity to interact with his lab group at Washington State University and provided field equipment and support during the early stages of my research. I thoroughly enjoyed our many conversations about range limits and salamanders. I had several great field assistants: Moses Michelsohn, Meghan Cooling, Nina Lobo, Stephen DeLisle, Matt Robinson, Talia Sechley and Chris Jacobs. Their help as part of Team Salamander was invaluable and I will always remember the fun times we had in the field together. Chris and Talia also helped with much of the animal care and generated two datasets for me.

An amazing group of friends got me through the many ups and downs of my PhD. Words fail to express my gratitude for love and companionship I found in Matt Siegle, Gina Conte, Anne Dalziel and Aleeza Gerstein. I also want to thank Gwylim Blackburn for being my “thesis buddy” during the final stretch; Milic Mandic for her kind and encouraging words and for several fun study dates; John Yensen for Thursday night hang-outs; and Rich FitzJohn for being my climbing partner for the better part of my time at UBC. For many other fun memories, I would like to thank Erika Crispo, Kira Delmore, Florence Debarre, Rowan Barrett, Alistair Blachford, Andrea Stephens, Jon Mee, Greg Sharam, Les Harris, Crispin Jordan, J.S. Moore, Sam Yeaman, Leithen M’Gonigle, Joel Heath, Carla Crossman, Jasmine Ono, Brook Moyers, Elizabeth Kleynhans, Bridgette Clarkston, Robin LeCraw, Angie Nichols, Heather Kharouba, Kate Ostevik, Kieran Samuk, Alana Schick, Kathryn Turner, David Toews, Rod Docking, Kay Hodgins, Kristin Ikswokrun, Amy Saini, Janice Domingo, Jeff Houde, Christine Grossen and Thor Veen.

For being “my person”, my light and my partner in the museum of ugly art project (amongst other fun things), I thank Stephen Jacob. Co-stalking the ginger cat, Sedona, creating the patio garden with the inadvertent Quebec-flag theme—our adventures, big and small, have meant the world to me these last couple of years.

Finally, I would like to thank my family: Bernadette and Jim Lee-Yaw, Jeff and Leah and the entire McEvoy clan. Their love, encouragement, guidance and support over the years got me to the finish line and continues to carry me forward.

This thesis is dedicated to my parents, Bernadette and Jim Lee-Yaw.

And to the salamanders, who still guard their secrets

Chapter 1: Introduction

1.1 Motivation

Explaining species' geographic range limits lies at the heart of understanding patterns of biodiversity. This challenge has long captured the attention of ecologists and evolutionary biologists alike (*e.g*. Darwin 1859; Mayr 1963; MacArthur 1972). Indeed, the study of species' geographic range limits requires the joint consideration of a number of ecological and evolutionary processes including dispersal and gene flow, adaptation and the evolution of the niche, as well as competition and other interspecific interactions (see reviews by Gaston 2003; Holt and Keitt 2005; Bridle and Vines 2007; Gaston 2009; Sexton *et al*. 2009; Wiens 2011). The need to evaluate these processes at multiple levels of biological organization (from genes to individuals to populations; Gaston 2009) makes the study of species' range limits a truly integrative line of research and a cornerstone in organismal biology.

However, the diversity of perspectives for approaching the problem of range limits also signifies the complexity of the problem. Consequently, most species' range limits remain poorly understood (Gaston 2009). Yet this information is precisely what is needed as we strive to predict and address the threats posed to species by environmental change. This need is the motivating force behind my research. My thesis specifically focuses on evaluating alternative hypotheses for parapatric range limits in a northern amphibian. In this first chapter, I briefly introduce parapatric range limits and review different hypotheses for species' range limits in the context of this distributional pattern. I then outline several questions that can help investigators better understand the relative importance of these hypotheses when studying parapatric range limits and discuss my approach to addressing these questions in the long-toed salamander (*Ambystoma macrodactylum*).

1.2 The case for studying parapatric range limits

Parapatry is a distributional pattern whereby two taxa occupy adjacent areas, sharing a narrow zone of overlap as a common boundary (Key 1982). The study of parapatric boundaries has a wide range of implications in ecology and evolutionary biology. For instance, parapatry has long been discussed in the speciation literature (*e.g*. Coyne and Orr

2004), particularly in the context of hybrid zones and the study of reproductive isolation (Hewitt 1988). However, many species exist in parapatry without hybridizing (*i.e.* "ecological parapatry" Bull 1991). Such cases provide basic information on the nature of competition or other factors preventing coexistence (Bull 1991). With respect to understanding large-scale patterns of biodiversity, parapatric range limits represent an opportunity to specifically evaluate the relative importance of biotic interactions versus other range-limiting factors in shaping species' distributions. Thus parapatric borders are natural testing grounds for a variety of questions pertaining to the maintenance and distribution of biodiversity.

Parapatric range limits are also ubiquitous. For instance, Hewitt (1989) estimates that between one-third and one-half of all sister species form hybrid zones on the landscape. This range of values almost certainly underestimates the number of parapatric boundaries given both the difficulties of identifying cryptic boundaries between morphologically similar taxa (Bickford *et al*. 2006) as well as the many cases of parapatry without hybridization. The prevalence of parapatric boundaries argues for their importance in shaping patterns of biodiversity. This point is especially true of northern areas where many species and populations have only recently (~10,000 ybp) come into (secondary) contact following isolation in separate glacial refugia (*e.g*. Hewitt 1996; Hewitt 2004).

1.3 Ecological and evolutionary explanations for range limits

The challenge that parapatric distributions present is explaining the factors that prevent the expansion of either species' range and thus range overlap (Bull 1991). Here, it is important to recognize that the factors limiting range expansion may differ for the two species involved in a parapatric border (*e.g*. Arif *et al.* 2007). Thus explaining parapatric distributions requires consideration of the mechanisms that can explain geographic distributions in the case of a single species as well as the potential role that interactions between species can have on a shared range limit. In this section, I discuss major ecological and evolutionary explanations for species' range limits, with an emphasis on the relevance of these hypotheses to parapatric distributions.

1.3.1 Range limits and the (fundamental) ecological niche

The concept of the ecological niche is highly relevant when discussing geographic distributions. For purposes of this thesis, I refer to the ecological niche of a species' in the Grinnellian sense, focusing on the effects of the environment on the species (rather than the Eltonian perspective of the role of the species in its community, as reviewed by Chase and Leibold 2003). Specifically, I use the term "niche" to refer to the set of resources and abiotic conditions required by a species for it to maintain non-negative population growth rates (*e.g.* Hutchison 1957).

To the extent that the requirements and tolerances of a focal species are known, the region of geographic space where it is possible for that species to persist at a given point in time can be defined (Peterson *et al.* 2011). This region is the spatial manifestation of the species' fundamental niche (*sensu* Hutchison 1957) and serves as a useful starting point for discussing range limits. For instance, the range limits of some species may simply reflect the limits of their physiological tolerances to abiotic conditions (*e.g.* temperature) or the distribution of a critical food item. However, range limits often fall short of the potential distribution of a species based on its fundamental niche (Jackson and Overpeck 2000; Svenning and Skov 2004; Araújo and Pearson 2005; Graham *et al.* 2010). Ecologists refer to that portion of the fundamental niche actually occupied by a species as its realized niche (*sensu* Hutchison 1957) and one question that arises is what leads to the discrepancy between observed range limits and the spatial extent of a species' fundamental niche? In the case of parapatric taxa, what factors preclude the expansion of one or both species when suitable conditions exist beyond the shared range limit?

1.3.2 Why do species' fail to fill their potential range on the landscape?

Insufficient time for colonization. One explanation as to why the range limits of some species do not reflect the distribution of suitable habitat is that insufficient time has passed for colonization (Paul *et al.* 2009; Wiens 2011). This explanation may be particularly relevant in northern areas as much of the northern hemisphere was covered by ice during the last glacial maximum. Temperate species have had ~10,000 years to recolonize these areas from southern refugia. The ranges of many of these species may still be expanding (*e.g.* Johnstone and Chapin 2003; Araújo and Pearson 2005) and thus some range limits may not

be at equilibrium. This explanation raises the possibility that some apparent cases of parapatry are simply temporary patterns on the landscape—a snapshot of two independent taxa in the midst of expanding their respective distributions or the initial phase of the invasion and take-over of one species' range by another (*e.g.* Bull 1991).

Dispersal Barriers. Range limits may settle at dispersal barriers. Dispersal barriers are organism-specific (Wiens 2011) and can be broadly defined as landscape features that restrict the movement and/or survival of individuals (see Stevens and Coulon 2012 for further discussion). The influence of dispersal barriers on species' geographic distributions is readily apparent when one considers their role in shaping patterns of colonization and the major reorganization of diversity that took place following the Pleistocene glaciations (*e.g*. Hewitt 1999). Some barriers preclude the movement of individuals and range expansion outright—an extreme example being the ocean limiting the ranges of terrestrial species. In other cases, dispersal barriers are regions where conditions are largely unsuitable for the long-term survival of individuals. Low levels of population connectivity in these regions may contribute to Allee effects (Keitt *et al.* 2001) or limit the recolonization of populations following local extinction (*i.e*. from a metapopulation perspective: Holt and Keitt 2000), both of which may lower population persistence and hinder subsequent range expansion. Recent studies have explored the role of dispersal barriers in maintaining parapatric distributions, demonstrating the potential for suitable habitat to exist for both species on either side of a common barrier to dispersal (Rissler and Apodaca 2007; Glor and Warren 2010; Soto-Centeno *et al.* 2013).

Biotic interactions. Species interactions can impact the extent to which a species fills its fundamental niche and are critical to consider in the case of parapatric distributions. For instance, in line with the competitive exclusion principal—which holds that ecologically identical species cannot coexist in space (Gause 1936)—several studies have found that competition shapes species' distributions at local spatial scales (*e.g*. Connell 1961; reviewed by Gaston 2003). Competition may similarly influence the broader range limits of some species, especially in the case of parapatric taxa (*e.g.* Bullock *et al.* 2000; Cunningham *et al.* 2009; Jankowski *et al*. 2010). Parasites and predators can also have demographic effects on populations that generate range limits under some conditions (Hochberg and Ives 1999; Holt and Barfield 2009).

Most models of the effects of species interactions on range limits address range limits that coincide with environmental gradients (*e.g.* MacArthur 1972; Case and Taper 2000; Holt and Barfield 2009). Species interactions may be less important determinants of those range limits that do not coincide with areas that already test the environmental tolerances of species (*e.g*. Case and Taper 2000). One exception that is highly relevant in the case of parapatric taxa is hybridization. Hybridization leading to the production of offspring that are intrinsically unfit can produce stable range limits, regardless of the ecological context of secondary contact (*e.g*. Anderson 1977; Barton and Hewitt 1985; Goldberg and Lande 2006). For instance, when hybrids are inviable or sterile, the rarer species is at a selective disadvantage and becomes rarer still through hybridization, creating a sharp parapatric boundary between species (Goldberg and Lande 2006). Even in cases where hybrid individuals are not completely unfit, selection against hybrids can maintain parapatric range limits, manifest as tension zones that reflect a balance between dispersal into the contact zone and selection against hybrids (Key, 1968; Barton and Hewitt 1985).

1.3.3 What limits niche expansion?

Parapatric range limits often coincide with ecotone boundaries or environmental gradients (reviewed by Bull 1991). In such cases, species may occupy different niches, with the shared range limit demarking the extent of suitable conditions for one or both species. For a species at the edge of its fundamental niche, range expansion requires adaptation to novel conditions—that is, niche expansion. The formation of stable range limits in these cases thus raises the broader question of what prevents adaptation at the periphery of a species' range (*e.g*. Haldane 1956; Mayr 1963; MacArthur 1972)? This question is not unique to parapatric range limits, and several hypotheses have be put forth to explain range limits that reflect a failure of adaptation. In addition to evaluating the relative importance of these hypotheses, parapatric boundaries challenge investigators to consider how interactions between the species themselves may limit niche expansion.

1.3.4 General explanations for the failure of adaptation at range limits.

Genetic Limitations. The evolution of some adaptations that would otherwise allow a species to expand its range may be unlikely, requiring either major mutations or

combinations of mutations (Hoffmann and Blows 1994). Local adaptation and range expansion may also be limited by low trait heritability at the range edge (Hoffman and Blows 1994). For instance, peripheral populations may have low levels of additive genetic variation owing to founder effects, genetic drift or strong directional selection (Eckert *et al.* 2008). Likewise, high levels of environmental variation may reduce the heritability of traits at the edge of species' ranges (Hoffmann and Blows 1994). Even when there is substantial genetic variation for a given trait, evolution may be limited by negative genetic correlations among traits or between different fitness components (Blows and Hoffman 2005). Thus several types of genetic constraints may preclude niche expansion and thus range expansion.

Gene flow. Mayr (1963) proposed that gene flow from populations at the centre of the range might stymie local adaptation at the periphery. Garcia-Ramos and Kirkpatrick (1997) modeled the evolution of a trait along an environmental gradient under the assumption of asymmetric gene flow from central to peripheral populations caused by a fixed density gradient (*i.e*. assuming an "abundant-centre" distribution: Sagarin and Gaines, 2002). In support of Mayr's hypothesis, they found that asymmetric gene flow towards the range edge can prevent peripheral populations from reaching local trait optima. Kirkpatrick and Barton (1997) extended this model, allowing population sizes to vary according to levels of local adaptation. Their results confirm that gene flow can cause trait means to deviate from optimal values as one moves away from the centre of the range and further demonstrate that the resulting decline in population fitness and size can lead to the formation of a range limit (Kirkpatrick and Barton 1997; see Filin *et al.* 2008 for exploration of how density dependence modifies these results). However, others have explored the positive effects of gene flow on local adaptation, finding that low to intermediate levels of gene flow can increase genetic variation in peripheral populations, thus facilitating range expansion (Barton 2001; Alleaume-Benharira *et al*. 2006). Gene flow between peripheral populations may specifically help promote the spread of beneficial alleles (Sexton *et al.* 2011). Thus too much or too little gene flow can potentially limit adaptation at the range margin and subsequent range expansion.

1.3.5 The potential for interactions between parapatric species to limit niche expansion

Competition. Competition can limit the evolution of traits that would permit range expansion. Case and Taper (2000) extended the models of Kirkpatrick and Barton (1997) to incorporate the effects of interspecific competition on adaptation along an environmental gradient. Their models demonstrate the potential for competition to reduce population size in areas of overlap, thus exacerbating the effects of asymmetric gene flow from central populations (Case and Taper 2000). Such gene flow may contribute to character displacement, offsetting competition to some extent, but ultimately causing populations to deviate from optimal trait values and leading to the formation of a parapatric range limit (Case and Taper 2000). Notably, when character displacement interacts with gene flow to limit local adaptation, stable range limits can be produced with less extreme environment gradients and levels of dispersal than those outlined by Kirkpatrick and Barton (Case and Taper 2000; see also Goldberg and Lande 2006). Even in the absence of gene flow, competition can prevent species from adapting to local conditions. Price and Kirkpatrick (2009) modeled the situation whereby two species utilize different resources. As one resource declines in space, the species specializing on that resource is prevented from adapting to the alternative resource by stabilizing selection arising from the effects of competition (Price and Kirkpatrick 2009).

Hybridization. As mentioned above, hybridization can lead to the production of stable range limits. The evolution of pre-mating barriers may be required before either species can expand its respective range. However, reproductive character displacement can lead to deviations from optimal trait values in sympatry, generating stable range limits (Goldberg and Lande 2006). These effects may be particularly pronounced along environmental gradients where reproductive character displacement can act in concert with gene flow from allopatric populations to limit the evolution of traits that would permit range expansion (Goldberg and Lande 2006).

1.4 Key questions for the study of parapatric range limits

The above discussion highlights the diversity of processes that can influence species' range limits and parapatric range limits in particular. Ultimately, a full explanation for a

given range limit requires evaluation of the relative importance of these hypotheses. Yet, comprehensive tests of even a single hypothesis are challenging, especially if the aim is to eventually have a truly mechanistic understanding of the traits and genes that govern range limits. Narrowing the list of hypotheses for direct testing is thus a critical phase of any research program aiming to understand range limits. In this section, I discuss several questions that are useful in this regard and that can serve to guide the investigation of parapatric range limits.

1.4.1 Where are the boundaries between distinct lineages?

1.4.1.1 Insight into factors influencing range limits

A first step to studying parapatric range limits is to accurately characterize the distributions of the focal taxa and locate shared borders. Apart from the obvious need to do so from the perspective of informing the specific location of subsequent study, the geographic context of parapatric boundaries may provide clues as to the mechanisms sustaining them and thus guide research efforts. For instance, a parapatric range limit that coincides with an obvious dispersal barrier in an otherwise uniform environment may argue for a different research focus than one that coincides with a sharp ecotone boundary.

1.4.1.2 The challenge of cryptic lineages

In practice, several logistical issues may complicate the task of locating range limits, including sparse locality records, issues with detection and the inaccessibility of some regions for direct survey. Locating parapatric range limits may require overcoming the additional difficulty of distinguishing between morphologically-similar taxa. Such cryptic taxa represent a particular challenge in northern regions where many nominal single "species" appear to be composites of highly distinct genetic groups—the legacy of repeated isolation in separate refugia during the Pleistocene glaciations (Hewitt 1996, 1999; Shafer *et al*. 2010). Identifying these cryptic lineages, evaluating the extent of divergence between them and delineating their boundaries not only defines the opportunity for the study of parapatric range limits in northern areas (*e.g*. Hewitt 1988) but also has implications for studying species' distributions more generally. For instance, depending on the treatment of cryptic lineages, perspectives may vary as to which populations are considered "central"

versus "peripheral" or whether species are considered to be "generalists" or "specialists" (*e.g.* Loxdale *et al.* 2011)—key classifications in the context of range limits. Such considerations underscore the need to have an accurate inventory of diversity in a focal system prior to undertaking studies of range limits.

1.4.2 What role does niche conservatism play in parapatric range limits?

1.4.2.1 Niche conservatism versus divergence and the biogeography of parapatry

Niche conservatism is broadly defined as the retention of ecological traits over time (Wiens *et al.* 2010). Across large time scales, niche conservatism manifests as phylogenetic niche conservatism—or the tendency for related species to demonstrate similarity in ecological traits (Wiens *et al.* 2010; see Losos 2008 for a more restrictive definition). Niche conservatism at this level of biological organization has been widely discussed in the context of understanding diversification, specifically as alternative hypothesis to ecological speciation (Wiens 2004; Kozak and Wiens 2006; Hua and Wiens 2013). However, the question of whether closely related taxa demonstrate niche conservatism (or conversely, niche divergence) also has implications for studying the maintenance of parapatric range limits. For instance, knowledge of the extent of ecological divergence between lineages, along with information as to whether a shared range limit coincides with the fundamental niche of either lineage, speaks to the distinction between hypotheses made earlier. Specifically, a range limit between ecologically similar taxa, with suitable habitat for both lineages on either side of the shared boundary, argues for investigation into dispersal limitation or the direct effects of species' interactions in setting range limits. In contrast, a range limit between two ecologically divergent taxa that coincides with the limits of suitable habitat for at least one lineage argues for investigation into those hypotheses that concern the failure of species to adapt to novel conditions or resources. This latter scenario is related to the concept of niche conservatism at a different level of biological organization—in that the evolution of ecological traits is constrained *within* species (Wiens *et al*. 2010).

Of course a failure to evolve wings or longer legs to overcome a dispersal barrier can still be considered a failure to adapt to conditions at the range edge. Thus even when range limits do not coincide with the extent of suitable conditions for a species, understanding the factors precluding the evolution of traits at range edges is important. However, determining

whether species are limited by the availability of suitable habitat versus dispersal barriers or biotic interactions can help investigators hone in on the most relevant traits for subsequent study. Even seemingly similar explanations for range limits at first glance may require different lines of investigation at the trait level. For instance, dispersal barriers may simply be localized regions of unsuitable habitat (*e.g*. Glor and Warren 2010) and thus related to the failure of individuals to adapt to novel conditions (Wiens 2011). However, the ability to disperse through a given environment may require the evolution of a very different set of traits (*e.g*. behavioral traits that govern the willingness of individuals to enter the environment, locomotive traits that allow individuals to move through the environment, etc.; see Figure 1 of Baguette *et al.* 2013) than would be necessary for the long-term survival and reproduction of individuals in the same environment. Thus evaluating the extent to which parapatric distributions are shaped by niche conservatism (both between and within taxa) lays the groundwork for more detailed examination of the traits and processes governing the shared range limit.

1.4.2.2 The role of ecology in diversification and the nature of species' interactions

The question of whether closely related taxa demonstrate niche conservatism (or conversely, niche divergence) also has implications for understanding the nature of species' interactions—a defining consideration in the case of parapatric range limits. For instance, the strength of competition—and thus the importance of this type of interaction for range limits—is expected to vary depending on the degree of ecological similarly between taxa (*e.g*. Letcher *et al*. 1994; Davies *et al*. 2007; Pigot and Tobias 2013). Likewise, in the case of hybrid zones, the extent of ecological divergence between taxa speaks to the potential importance of exogenous reproductive barriers for the maintenance of genetic boundaries. Thus evaluating whether closely related taxa demonstrate niche conservatism (or conversely, niche divergence) not only speaks to the importance of ecology in promoting diversification in the first place, but can also inform predictions regarding the subsequent outcome of interactions at shared range limits.

1.4.3 What is the nature and outcome of interactions between lineages at a shared border?

1.4.3.1 Confirmation that lineages are interacting

Species' interactions are generally assumed to impose a constraint on range expansion in the case of parapatric taxa. However, several of the explanations listed in Section 1.3, highlight the potential for the ranges of the taxa involved to be relatively independent of each other. Thus directly demonstrating interactions between lineages is a necessary step when studying parapatric distributions. This need is particularly important given the potential for taxa to be interacting in unexpected ways. For instance, competition is widely regarded as the most important interaction influencing range limits between non-hybridizing lineages (Bull 1991). However, direct demonstration of competitive interactions (*e.g*. Cunningham *et al*. 2009) is necessary to rule out alternatives, including the effects of shared parasites or predators (*i.e*. "apparent competition": Dawes-Gromadzki and Bull 1997; Ricklefs 2010). Likewise, the potential for closely related forms to demonstrate very little, if any, hybridization in secondary contact (*e.g.* Irwin *et al.* 2001) cautions against *a priori* assumptions that reproductive interactions drive a shared range limit.

1.4.3.2 Linking process to population demography

Gaston (2009) presents the case for linking specific range-limit hypotheses to birth and immigration rates versus death and emigration rates in peripheral populations. Beyond demonstrating that a particular interaction is occurring at a parapatric boundary, it is thus necessary to address the consequences of the interaction for individual and population fitness. For instance, most hypotheses invoking hybridization at range limits assume that hybrid zones represent demographic sinks, imposed by the low fitness of hybrid individuals. However, although natural hybridization is widely documented, comparably few studies have directly assessed the fitness of hybrids, especially in northern contact zones (see Kruuk *et al.* 1999; Turner *et al*. 2011; Ålund *et al*. 2013 for examples). In light of the potential for there to be other outcomes of hybridization upon secondary contact (*e.g.* hybrid swarms: Wiens *et al.* 2006, bounded hybrid superiority: Moore 1977), directly evaluating the fitness of hybrids is a critical step when studying the importance of hybridization for range limits (see also Arnold

1995). A parallel case can be made for assessing the relative fitness of individuals experiencing competition a parapatric boundary (see Cunningham *et al.* 2009 for example).

1.5 The Long-Toed Salamander as a study system

The long-toed salamander is a promising system for investigating parapatric range limits. This species is one of the most widely distributed amphibians in western North America. Within its range, five subspecies are currently recognized, described on the basis of variation in morphological trait values across the range (Ferguson 1961) as well as a handful of genetic markers (Thompson and Russell 2005; Savage 2008). Although additional genetic work is necessary to clarify the extent to which these subspecies represent distinct evolutionary lineages and to accurately delineate their ranges, current descriptions of the subspecies' distributions suggest that they form several, highly elongated (mainly running from north to south) parapatric boundaries in the Pacific Northwest. Thus the system potentially affords an opportunity to take a comparative approach to the study of parapatric range limits—allowing for identification of the factors contributing to range limits at multiple locations within the same parapatric boundary as well as across the parapatric boundaries of different pairs of lineages.

Several aspects of the geography of the region and ecology of the long-toed salamander suggest that the system also represents an excellent opportunity to explore the relative importance of different hypotheses for parapatric range limits. Clearly if areas of sympatry between distinct genetic groups are identified, then the boundaries between subspecies are ripe for exploring the outcome of secondary contact and the role that hybridization and competition play in generating parapatric range limits. In addition, the long-toed salamander is found in one of the most topographically complex regions of North America. In particular, the Pacific Northwest encompasses several mountain ranges and major rivers. A number of landscape genetic studies have found that these landscape features influence long-toed salamander dispersal at fine-spatial scales (Tallmon and Funk 2000; Giordano *et al*. 2007; Savage *et al*. 2010). Dispersal barriers might thus be expected to play a major role in shaping broad-scale patterns of diversity in this system.

The Pacific Northwest is also diverse with respect to climatic conditions. This environmental variation has promoted narrow and/or disjunct distributions in other species in

the region (Brunsfeld *et al.* 2001) and may similarly influence the distribution of major phylogenetic groups in the long-toed salamander. Certainly climate does have an effect on variation among long-toed salamander populations at local scales. For instance, adult individuals from different temperature regimes are known to demonstrate differences in critical thermal maxima (Howard *et al.* 1983). Temperature tolerance during development is similarly known to vary among larvae from some populations (Anderson 1972). Likewise life-history traits (clutch and egg size; time to and size at metamorphosis) are dramatically different between populations at different elevations (Howard and Wallace 1985; Anderson 1967). At least some of this variation appears to have a genetic basis as a recent study found that life-history traits differed between larvae collected at different elevations and reared under common conditions (it is noted that maternal effects cannot be ruled out in this experiment; Giordano 2005). Although many of these differences occur on small spatial scales and may reflect local adaptation within lineages, some authors have suggested that such variation may play out across subspecies boundaries (Anderson 1967). Thus the role of climate in shaping patterns of diversity in this system warrants investigation. Overall, the multiple candidate factors influencing range limits in the long-toed salamander underscores the utility of the system for exploring the relative importance of different hypotheses for parapatric range limits.

1.6 Specific objectives and overview of thesis

The overall goal of my research was to characterize and evaluate the relative importance of different factors in shaping the distribution of diversity in the long-toed salamander. In line with the first question from the above section on studying parapatric range limits, my first objective was to describe phylogenetic diversity and document range limits in this system. Using multilocus genetic data, I evaluated the extent to which currently-recognized subspecies represent distinct evolutionary units, testing the hypothesis that the large geographic range of the species is actually a composite of the smaller ranges of several, highly distinct genetic groups (Chapter 2). My work demonstrates a novel approach for using genomic data in combination with spatial analysis to delineate parapatric range limits (Chapter 3). I note that in comparing single-gene phylogenies to a genome-wide assessment of genetic structure (Chapter 2), my results also speak to the question of whether the

boundaries between closely related taxa can be adequately characterized by just a handful of markers.

With the diversity and distributions of long-toed salamander lineages thus described, I turned to the task of evaluating the extent of niche divergence between lineages and the correspondence between range limits and the fundamental niches of the different lineages (Chapter 3). Given its influence on broad-scale patterns of biodiversity in the Pacific Northwest (Brunsfeld *et al.* 2001) and on variation among long-toed salamander populations (see Section 1.5), I focused on the role of climatic in shaping lineage boundaries. For each parapatric pair of lineages, I specifically addressed two hypotheses. The first maintains that each lineage is uniquely adapted to its existing range, with range limits simply delineating the extent of their respective climatic niches (*e.g*. Rissler and Apodaca 2007; Glor and Warren 2008). The second holds that ecologically similar taxa occupy either side of a climatic barrier—a region of low climatic suitability separating otherwise suitable habitat for both lineages (Glor and Warren 2008). My work brings together several different frameworks for testing these hypotheses with spatial data and ecological niche models (*e.g.* Rissler and Apodaca 2007; Glor and Warren 2010; McCormack *et al.* 2010; Arteaga *et al*. 2011) and I note that my results provide one of the few tests of variation in the extent to which climate influences a given parapatric boundary along its length (see also Werner *et al.* 2013).

Finally, I explored the role of hybridization in shaping lineage boundaries. Results from my genome-wide survey of genetic variation suggest that the lineages hybridize where they come into contact (Chapter 2). To further characterize the extent of hybridization, I conducted a fine-scale survey of genetic variation across a sharp mitochondrial divide between two of the lineages in southern British Columbia and Alberta (Chapter 4). To better understand the consequences of hybridization, I explored the potential for there to be hybrid dysfunction in the contact zone (Chapter 4). Tests of niche conservatism (Chapter 3) along with previous field observations suggest most long-toed salamander lineages are ecologically similar and thus I focused on evaluating the potential for intrinsic genetic incompatibilities between lineages to be shaping the outcome of secondary contact. Specifically, I asked whether the feeding performance and energy conversion efficiency of individuals varies with respect to their genetic ancestry. I note that this metric of organismal performance, though

widely studied in other regards (*e.g*. Imsland *et al.* 2001; Grayson *et al.* 2005; Baumann and Conover 2011), has not been previously considered with respect to hybridization and reproductive isolation. Thus my work extends measures of hybrid dysfunction to include an important but currently understudied aspect of individual vigor.

In summary, my dissertation explores the following questions with respect to understanding parapatric range limits in the long-toed salamander:

1. How is genetic diversity partitioned in the long-toed salamander? (Chapters 2 & 3)
2. To what extent do the different lineages of long-toed salamander demonstrate niche conservatism versus niche divergence? (Chapter 3)
3. Do the range limits of the different long-toed salamander lineages reflect the limits of their respective climatic niches? (Chapter 3)
4. Is there evidence for hybridization (Chapters 2 & 4) and hybrid dsyfunction (Chapter 4) in contact zones between long-toed salamander lineages?

Chapter 2: Large geographic range size reflects a patchwork of divergent lineages in the Long-Toed Salamander (*Ambystoma macrodactylum*)

2.1 Summary

For northern taxa, persistence in multiple versus single Pleistocene refugia may have been an important determinant of contemporary range size, with larger ranges achieved by species that colonized the north from several glacial refugia. Under this hypothesis, widespread species are expected to demonstrate marked phylogeographic structure in previously glaciated regions. In this study I use a genome-wide survey to characterize genetic structure and evaluate this hypothesis in the most widely distributed salamander in the Pacific Northwest, the long-toed salamander (Ambystoma macrodactylum). Patterns of variation based on 751 AFLP loci and mitochondrial sequence data were concordant and support the recognition of at least four distinct lineages of long-toed salamander. The distributions of these lineages indicate that multiple refugia contributed to the species' large contemporary range. At the same time, with up to 133 AFLP bands differing between lineages and levels of sequence divergence ranging from 2.5 to 5.8%, these lineages would be considered separate species by some definitions. Such splitting would partition the large geographic range of the long-toed salamander into several relatively restricted ranges. These results thus underscore the potential for estimates of geographic range size to vary considerably depending on the taxonomic treatment of cryptic lineages.

2.2 Introduction

That most species have relatively small geographic distributions is a well-established pattern in macroecology (reviewed by Gaston 2003). Identifying the factors that have permitted some species to achieve comparably large distributions is a major goal in ecology, with leading hypotheses invoking traits such as dispersal ability (Lester *et al*. 2007, Leger and Forister 2009), body size (Cambefort 1994; Gaston and Blackburn 1996) and niche breadth (Garcia-Barros and Romo Benito 2010). Recently, Shafer *et al.* (2010) demonstrated an association between contemporary range size and persistence in multiple versus single glacial refugia in western North America, adding a potentially important historical dimension

to understanding large geographic ranges in previously glaciated regions. Specifically, this pattern suggests that persistence in multiple glacial refugia and the patching together of the post-glacial ranges of several populations (Hewitt 2004) was instrumental to the formation of large ranges in northern areas.

Isolation and divergence in multiple glacial refugia during the Pleistocene has also been implicated in speciation in some taxa (Weir and Schluter 2004; Levsen *et al*. 2012; but see Barber and Jensen 2012 and references therein). Thus the pattern observed by Shafer *et al.* (2010) raises the additional possibility that some widespread species actually consist of genetic groups that, following taxonomic scrutiny, will warrant species-status. Corresponding reductions in described geographic ranges may have important implications for comparative studies that rely on estimates of range size to test hypotheses for variation in range size (see Loxdale *et al*. 2011 for related discussion).

Characterizing genetic structure in northern areas is an important first step towards determining whether widespread species represent patchworks of divergent groups and the extent to which such groups warrant taxonomic recognition. Although many existing phylogeographic studies speak to this goal, most have been conducted using just a handful of genetic markers. The potential for stochasticity and selection to cause patterns for any single marker to deviate from the overall history of closely related groups limits the conclusions that can be drawn from studies using a small number of markers (Irwin 2002; Rauch and Bar-Yam 2004; Dowling *et al.* 2008; Irwin 2012). Methods that survey many markers throughout the genome overcome this concern and inherently assess the degree of divergence between taxa. Because the use of a limited number of genetic markers often leads to the underestimation of genetic diversity (*e.g.* Irwin *et al*. 2009; Mila *et al*. 2010; Wiens *et al*. 2010), revisiting phylogeographic structure with these high-resolution methods may be particularly pertinent for determining whether widespread species represent single cohesive genetic entities or composites of genetically distinct groups.

The long-toed salamander (*Ambystoma macrodactylum* Baird, 1849) exemplifies the case of a widespread species for which phylogeographic patterns and the extent of diversification remain unclear, despite previous attention. Currently considered one species, the long-toed salamander is the most widely distributed salamander in the Pacific Northwest, a region that is both geologically and ecologically complex (Brunsfeld *et al.* 2001). The

potential for the species' wide range to consist of multiple distinct lineages was first suggested by Mittleman (1948) and later Ferguson (1961), whose work examining vomerine tooth count, allometry and strip pattern and colouration in the southern portion of the species' range led to the recognition of five subspecies of long-toed salamander (Ferguson 1961; Figure 2.1). More recently, two studies employing a limited number of genetic markers (mtDNA: Thompson and Russell 2005; mtDNA and six nuclear loci: Savage 2008) have provided some evidence of genetic differences between named subspecies. However, different patterns observed among these studies and among markers within the study by Savage (2008), as well as discrepancies between genetic data and morphological subspecies (Thompson and Russell 2005; Savage 2008) make the interpretation of subspecies difficult. Furthermore, limited sampling to date, especially in northern areas, restricts the conclusions that can be drawn about the distribution of distinct lineages and the phylogeographic structure of the species in previously glaciated regions. The extent to which previously described groups are representative of genome-wide patterns of differentiation thus requires further assessment.

In this study, I undertake a comprehensive genetic survey to better understand diversification in the long-toed salamander. Comparing phylogeographic patterns observed using amplified fragment polymorphisms (AFLPs) to those identified by mitochondrial DNA, I ask whether genome-wide patterns of differentiation support the existence of multiple lineages of long-toed salamander. I simultaneously use the genetic data to determine the distribution of major genetic groups, with a focus on northern parts of the species' range. In doing so, I clarify the extent to which one of the largest geographic ranges in the Pacific Northwest represents a truly remarkable feat of colonization by a single taxonomic entity versus a patchwork of multiple divergent lineages.

2.3 Materials and methods

2.3.1 Sampling

Tail clips were collected during the spring and summers of 2008 through to 2010. Other researchers and the Museum of Vertebrate Zoology (University of California, Berkeley) provided additional samples. In total, 403 individuals from 122 sites across the range of the species were sampled (Figure 2.1; Table A1).

2.3.2 AFLP data collection

Total genomic DNA was extracted using a standard phenol-chloroform protocol. I generated AFLP profiles for 378 individuals from 108 sites (Figure 2.1). The 25 remaining individuals were only included in the mitochondrial dataset (below) due to either poor DNA quality or because they were obtained after the ALFP data had been generated. I used the AFLP protocol of Vos *et al*. (1995), modified according to Toews *et al.* (2008), for the digestion, ligation and pre-amplification steps. Selective amplification followed the protocol of Clarke and Meudt (http://clarkeresearch.org/aflp_2012-01-26/AFLP_Protocol.pdf). To reduce the complexity of banding patterns resulting from the large size of *Ambystoma* genomes, I included two selective amplification steps (Voss and Shaffer 1997). Using these protocols, I generated two AFLP datasets. The most inclusive dataset of 378 individuals was generated using one selective primer combination in the final selective amplification (Table 2.1). A subset of 70 of these individuals from 39 populations (Figure 2.1) were then screened for an additional five selective primer combinations. This second dataset thus represents more comprehensive sampling of the genome for a representative set of individuals.

Fragment detection was conducted on an automated ABI 3100. AFLP electropherographs from each primer combination were imported into Peak Scanner v.1.0 (Applied Biosystems 2006) for initial analysis. Peak sizes were called using default settings except for the application of lite smoothing recommended for automatic scoring (Arrigo *et al.* 2012). The automatic binning and scoring algorithm implemented in RawGeno v. 2.11.1 was then used to analyze peaks exported from PeakScanner for each primer pair. The bin detection range was 50 to 400 bp to minimize homoplasy associated with very small fragments (Vekemans *et al*. 2002) and detection issues associated with drop-off at larger fragment sizes. Individuals were scored as having a band "present" for a bin if a corresponding peak exceeded 80 rfu. To reduce the number of uninformative bins, I eliminated bins for which fewer than five individuals were scored as having the band present (or absent). Additionally, 5 to 13% of individuals (depending on primer pair), replicated from DNA extraction, were included in the set of samples for each primer pair. I set the repeatability filter in RawGeno to remove bins that were <80% repeatable across these replicates (cutoff recommended by Nils Arrigo pers. com.). After filtering bins, the two

AFLP datasets included 177 loci (average repeatability 92%) and 751 loci (average repeatability of 91.3%) respectively.

2.3.3 Mitochondrial DNA data collection

For comparison with the AFLP data, I sequenced the mtDNA cytochrome b gene from 142 individuals. Primers were designed from conserved regions of the mitochondrial genomes of *A. laterale* (GENBank Accession NC_006330; Mueller *et al*. 2004), *A. mexicanum* (GENBank Accession NC_005797; Arnason *et al.* 2004) and *Plethodon petraeus* (GENBank Accession NC_006334; Mueller *et al*. 2004): AmbPleth_cytb-F 5'ACYGRAACCYTTGACMTGAA; AmbPleth_cytb-R 5'YCRRTTTTCGRCTTACAAGG. Sequences were obtained using nested primers for two museum specimens from the Sierra Nevada Mountains and Santa Cruz (California) that were of reduced quality and thus not included in the AFLP dataset. PCR was carried out in 25-μl reactions consisting of 0.25 μl dNTPs (10 μM), 2.5 μl 10x reaction buffer (Invitrogen), 0.75 μl MgCl (50 mM), 1 μl each primer (10 μM), 18.4 μl ddH_20, 1 μl TAQ (5000 U/ml: New England Biolabs) and 1 μl DNA (25 ng/μl). Amplification involved initial denaturation at 94°C for 2 minutes, 35 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 45 s and extension at 72°C for 1 minute and 10 s, and a final extension at 72°C for seven minutes. PCR products were sequenced by the Genome Quebec Innovation Centre at McGill University on an automated ABI 3730XL. The resulting chromatographs were verified by eye. Ten individuals (including the two California samples) were initially sequenced in both directions and compared in BioEdit Version 7.0.5.3 (Hall 1999). As all sequences were unambiguous and identical in both directions, I sequenced the remaining individuals using the reverse primer only. Use of this primer resulted in 729 bp of sequence data. Sequences were manually verified and edited in BioEdit v. 7.0.5.3 (Hall 1999) and aligned using Clustal (Thompson *et al.* 1994).

Based on analysis of the sequence data (below), I obtained mtDNA group membership information for the remaining 261 individuals included in the AFLP dataset using PCR-RFLP (Appendix A2). Briefly, samples were amplified for a 597 bp section of the *cyt b* gene described above and then digested with up to three restriction enzymes that allowed for the unambiguously assignment of individuals to mitochondrial group based on fixed SNPs observed in the sequence data. Thus mtDNA data was available for all sites

included in the study, allowing me to compare mtDNA lineage boundaries with those observed for the nuclear genome.

2.3.4 Evaluating genetic structure

I used two methods to determine whether genome-wide patterns revealed by the AFLP data support distinct lineages of long-toed salamander. For both AFLP datasets, the Bayesian clustering algorithm implemented in STRUCTURE (Pritchard *et al.* 2000; Falush *et al*. 2007) was used to calculate posterior probabilities of the assignment of individuals to 1 to 12 groups (K) using the admixture model of ancestry. For each K, I conducted 10 runs of 1000000 MCMC generations with the first 500000 generations discarded as burn-in. The optimal value of K was determined by examining the log probability of the data as well as using the method outlined by Evanno *et al*. (2005). For the AFLP dataset generated from six selective primer combinations (*i.e.* the largest survey of the genome), I also examined similarity between individuals using a multidimensional scaling analysis based on Jaccard distances calculated using the Vegan package in R (Oksanen *et al*. 2011).

Relationships amongst individuals based on mtDNA were explored using the haplotype network estimation procedure implemented in TCS version 1.21 (Clement *et al*. 2000). Due to the sensitivity of the analysis to ambiguous data, I excluded characters for which any individual had ambiguous base calls. Network estimation was performed on two data partitions. The first partition of 600 unambiguous characters excluded the sole representative from the Sierra Nevada Mountains in California as I had limited sequence data for this individual and wished to infer relationships between haplotypes with the largest number of characters possible. The second partition was a 489 unambiguous character subset of the data that allowed inclusion of this individual for purposes of determining its likely position in the network.

2.3.5 Assessing degree of divergence

I used the 70 individuals for which I had the largest amount of AFLP data to assess genetic divergence between the major genetic groups. Arlequin 3.11 (Excoffier et *al*. 2005) was used to calculate the average distance between groups in terms of the number of bands and to perform an analysis of molecular variance (AMOVA) to evaluate the percentage of

AFLP variation attributed to differences within versus between major groups. These calculations required *a priori* assignment of individuals to groups. In order to avoid circularity, I assigned individuals to groups based on the major mitochondrial breaks suggested by the haplotype network rather than using the nuclear DNA groups suggested by STRUCTURE. I also calculated average uncorrected and corrected pairwise sequence divergence between the mitochondrial groups using the APE package in R (Paradis *et al.* 2004). Corrected differences were calculated using the Tamura and Nei (1993) model of sequence evolution with gamma parameter.

2.4 Results

2.4.1 Genetic structure

Substantial genetic structure was observed in the AFLP datasets. Examination of mean ln probabilities and calculation of ΔK (Evanno *et al*., 2005) following the STRUCTURE analyses revealed a clear peak at K=5 for both datasets (Figure A1). Four of these groups are geographically distinct, partitioning the species into a Coastal-Cascade group, a North-Central group, a Rocky Mountains group and a Central Oregon Highlands group (Figure 2.2). The proportion of each individual's ancestry from the fifth group was generally low (although there were a few exceptions in the analysis of 177 loci: Figure 2.2). There was no clear geographic structure associated with this fifth group.

The multidimensional scaling analysis of 751 AFLP loci resolved three of the genetic groups observed in the STRUCTURE analysis: the Coastal-Cascade, North-Central and Rocky Mountains groups (Figure 2.3). Individuals from the Central Oregon Highlands clustered with Coastal-Cascade individuals, although there was slight divergence between these groups along Dimension 1 of the MDS scaling plot. Close association between individuals from the Central Oregon Highlands and Coastal-Cascade Mountains is also suggested by the clustering of these individuals as one group in the STRUCTURE analysis of k=4, one less than the optimal value (plot not shown).

Strong phylogeographic structure was also evident in the mitochondrial genome. Specifically, haplotype network estimation using the dataset with the larger number of unambiguous characters resulted in six distinct networks that could not be connected within the limits of statistical parsimony, in this case 10 mutational steps (Figure 2.4). Four of these

networks coincide with the groups resolved by the STRUCTURE analysis (Coastal-Cascade, North-Central, Rocky Mountain and Central Oregon Highlands groups), although several individuals from populations close to the geographic transition between AFLP groups demonstrated a mismatch between mtDNA group and AFLP cluster membership (Figures 2.2 and 2.5). The sole sample from the currently described Santa Cruz subspecies came out as a separate network, as did two individuals from Southwestern Oregon.

Haplotype network estimation with fewer characters (to permit the inclusion of an individual from the Sierras in California), resulted in similar relationships to the analysis with the larger character set. However, individuals from Southwestern Oregon were distantly connected to the Coastal-Cascade network rather than representing a distinct network (Figure 2.4). The haplotype from the Sierras also connected to the Coastal-Cascade network with a large number of mutational steps (Figure 2.4). Both iterations of the analysis suggest divergence between northern and southern haplotypes within networks (Figure 2.4).

2.4.2 Degree of divergence

From a survey of 751 AFLP loci from across the genome, the average number of AFLP bands that differed between pairs of lineages ranged from 90-133 (Table 2.2). Analysis of molecular variance (AMOVA) suggests that a significant proportion of variation in the AFLP data is attributed to differences between lineages (Percent Variation = 14.9, d.f. = 3, SSD = 570.33, $p << 0.05$); although most of the variation observed was due to differences within lineages (Percent Variation = 85.0, d.f. = 65, SSD = 3281.11, $p << 0.05$). Average uncorrected pairwise sequence divergence between mitochondrial groups ranged from 2.4% to 5.2% (2.5% to 5.8% when corrected according to Tamura and Nei [1993]; Table 2.3). Sequence divergence within groups was substantially lower ranging from 0.06% to 1.8%.

2.5 Discussion

Results from my genome-wide survey of AFLPs and from the mitochondrial dataset are highly concordant and clearly support the existence of multiple lineages of long-toed salamander. The distributions of these lineages and degree of divergence among them indicate that several refugial populations contributed to the large contemporary range of the

species. Furthermore, the relatively restricted geographic areas occupied by the individual long-toed salamander lineages highlight the potential for there to be large discrepancies between estimates of geographic range size depending on the taxonomic treatment of cryptic lineages.

2.5.1 Diversity within the Long-Toed Salamander

My analysis of several hundred AFLP loci and mtDNA sequence data support at least four major lineages of long-toed salamander. These results thus complement previous studies (Mittleman 1948; Ferguson 1961; Thompson 2003; Thompson and Russell 2005; Savage 2008), providing genome-wide evidence of genetic differences between some of the described subspecies. These results also contribute much needed clarification of the boundaries of these and other cryptic groups within the species, highlighting several key places where described subspecies' boundaries are inaccurate or do not adequately capture diversity within the species (see Figure 2.5 for summary).

One of the most notable results from my genetic survey was observation of a cryptic lineage in the southern portion of the currently described range of *A. m. columbianum* (Figure 2.5). Individuals from this region come out as a distinct group in STRUCTURE analyses of K=5 (the optimal K) and harbour unique mtDNA haplotypes. These individuals cluster closely with Coastal-Cascade individuals in the MDS analysis (at least in the two dimensions considered presently) and when the STRUCTURE analysis is run at K=4. Intriguingly, however, mtDNA haplotypes from this region appear most closely related to haplotypes from the North-Central portion of the species' range. When the number of mutational steps permitted in the haplotype network analyses is increased to 12 (beyond the limits of statistical parsimony) haplotypes from the Central Oregon Highlands form a distantly related group within the North-Central haplotype network. Pairwise sequence divergence between these two groups is also low (Table 2.3). Previous analyses of mtDNA data have placed two samples from central Oregon into a clade encompassing *A. m. columbianum* (Savage, 2008). Thus populations from the Central Oregon Highlands have an interesting history, with most analyses distinguishing these populations as a unique lineage but with different markers suggesting different histories of association with other lineages.

The mitochondrial data point to additional genetic structure across the species' range. Most notably, the haplotype from the sole individual representing the very spatially-restricted Santa Cruz isolate could not be connected to any other network within the limits of statistical parsimony in the haplotype network analysis. High levels of sequence divergence between this haplotype and haplotypes from other places (Table 2.3) suggest that the Santa Cruz region harbors a distinct group of long-toed salamanders (see also Savage 2008). Likewise, an individual from the Sierra Nevada Mountains, representing what is currently described as *A. m. sigillatum*, was highly distinct in the haploytpe network. However, the degraded quality of DNA from these two individuals precluded me from obtaining AFLP data to test whether these patterns are representative of the entire genome and more more extensive sampling of the southern-most portion of the species' range is required before such conclusions can be reached. Such rigorous assessment is particularly pertinent given several examples where genetic structure is apparent in the mitochondrial data but not in the AFLP data. For instance, samples from Southwestern Oregon come out as a distinct haplotype network (Figure 2.4) but have AFLP profiles that fall clearly within the Coastal-Cascade lineage. Likewise, none of the distinct groups within each major haplotype network (Figure 2.4) were observed in any of the AFLP analyses suggesting that much of the population structure observed in the mtDNA reflects more recent divergence.

2.5.2 Widespread species as a patchwork of the post-glacial ranges of many populations

Results from a recent study by Shafer *et al.* (2010), demonstrating a positive association between geographic range size and persistence in single versus multiple glacial refugia, highlight the potential for colonization from multiple refugia to have been important for the generation of large geographic ranges in northern areas. The long-toed salamander was one of the most widespread species featured in that analysis; yet conclusions about the number of refugial populations were based on a very limited dataset (288 bp of mtDNA sequence data) that supported only two lineages of long-toed salamander in previously glaciated areas (see Figure 2 of Thompson and Russell 2005). My genome-wide survey of diversity points to the existence of at least one additional lineage of long-toed salamander in the north (see also Savage 2008), lending even more support to the association between range

size and number of refugia observed by Shafer *et al*. (2010) and highlighting the potential for high-resolution markers to uncover diversity within widespread species that may have been missed by studies using a limited number of markers.

With respect to the history of the species' range, several lines of evidence indicate that the long-toed salamander maintained populations in multiple glacial refugia during the Pleistocene. Informal estimates of divergence using a rate of change of 0.7 to 1% sequence divergence per million years (based on molecular clocks calibrated using other salamander species as reviewed by Caccone *et al.* 1997) suggest that the split between the most closely related long-toed salamander lineages occurred at least 2.4 million years ago. Although informal, these estimates preclude divergence after the last glacial maximum (*i.e*. ~20,000 ybp) and thus rule out a single post-glacial colonization event followed by diversification. Likewise, the average number of AFLP bands (*i.e*. SNPs) that differed between lineages was high (Table 2.2), suggesting an extended history of population isolation. Necessarily, all of the lineages described here have clear overlap with many of the previously proposed glacial refugia in the Pacific Northwest (inset of Figure 2.1 and Table 2.4). Thus rather than representing an impressive feat of colonization of the Pacific Northwest by a single lineage, the large contemporary range of the long-toed salamander complex is best explained by the long-term persistence of at least four geographically separated populations during the Pleistocene glaciations and the collective post-glacial expansions of these individual populations. Such fusing of glacial isolates may have been critical in allowing some northern taxa to become widespread, especially species such as the long-toed salamander for which there is evidence of limited dispersal capabilities and a sensitivity to dispersal barriers (Tallmon *et al*. 2000; Giordano *et al.* 2007; Savage *et al*. 2010; Goldberg and Waits 2010) that might otherwise have hindered colonization of the full range.

2.5.3 Taxonomy and implications for geographic range size

Formal taxonomic evaluation of the long-toed salamander is beyond the scope of this study. However, I note that taxonomic treatment of the different lineages fundamentally affects estimates of geographic range size. As a single species, the described geographic range of the long-toed salamander is approximately 13×10^5 km^2 (based on the range map provided by the Global Amphibian Assessment Database: IUCN *et al*. 2008). In contrast, the

geographic range sizes of the individual long-toed salamander lineages are magnitudes lower, varying from approximately 730 (putative Santa Cruz lineage supported by the mtDNA) to 58 x 10^4 km^2 (North-Central). The latter estimates are well within the distribution of range sizes of other amphibians in the Pacific Northwest. Descriptions of the long-toed salamander as a 'widespread species' thus depend critically on taxonomic treatment of the different lineages.

Whether the different long-toed salamander lineages should be considered separate species depends on the species' concept employed (see Coyne and Orr 2004 for review of species' concepts). Genome-wide differences between the lineages and levels of mtDNA sequence divergence consistent with what has been reported among young amphibian sister species (giant salamanders in the genus *Dicamptodon*, 4.3 to 6.7%: Steele *et al.* 2005; torrent frogs in the genus *Amolops*, <1 to 3.1%: Matsui *et al.* 2006) would argue for splitting the species under many genetic definitions (although I note that others have reported similar levels of mtDNA variation between what are still considered subspecies within species [*e.g.* toads in the *Bufo americanus* complex, 1.8 to 3.96%: Masta *et al*. 2002]). Morphological differences (Ferguson, 1961) and putative variation in life-history (Anderson 1967; see also Kezer and Farner 1955; Howard and Wallace 1985) and diet (Anderson 1968) among some subspecies would also support the recognition of distinct species from an ecological standpoint. However, my data reveal several cases of cyto-nuclear discordance in the contact zones between lineages, indicating that some hybridization does occur. The extent of such cyto-nuclear discordance and hybridization is generally limited to narrow (50-125 km-wide) contact zones between lineages (see Figure 2.5). Nevertheless, fine-scale genetic sampling across the lineage boundaries identified presently, as well as experimental tests of reproductive isolation between lineages are necessary to further characterize these contact zones and determine whether sufficient reproductive isolation exists to consider these lineages species under the widely-employed biological species concept (Mayr 1963; Coyne and Orr 2004).

Although the taxonomic status of the different long-toed salamander lineages requires further evaluation, the finding that the most widely distributed salamander in the Pacific Northwest consists of distinct, relatively geographically restricted genetic groups underscores the value of phylogeographic assessment of species' range limits. Several other

phylogeographic studies reveal cases where cryptic genetic groups—should they warrant taxonomic recognition—necessitate significant reductions to the described ranges of widespread species (*e.g.* Irwin *et al*. 2001; Toews and Irwin 2008; Oliver *et al*. 2009; Tan *et al.* 2010). Such overestimates of species' ranges may be common if 'widespread species' in northern hemispheres generally do have a history of persistence in multiple glacial refugia and thus represent patchworks of potentially very divergent lineages (*e.g.* Berggren *et al*. 2005; Niedzialkowska *et al.* 2011; additional examples in Shafer *et al.* 2010). Identification and taxonomic scrutiny of genetic lineages within these species is pertinent for interpreting the results of the many studies that use published estimates of species' distributions (*e.g.* NatureServe) to explore hypotheses concerning geographic range size.

2.5.4 Outstanding questions

That some widespread species represent patchworks of distinct populations that have withstood the test of time raises a number of additional questions. Critically, post-glacial colonization from multiple refugia may explain the large contemporary distribution of some northern species such as the long-toed salamander; however, what allowed these species to initially maintain populations in widely separated refugia? Did species with currently restricted ranges that suggest a history of expansion from a single refugium once have wider ranges and simply suffer greater extinction during the Pleistocene glaciations? If so, what factors explain such variation in persistence during environmental change (*e.g*. Davies *et al.* 2009; Waldron 2010)? Additionally, are widespread species from northern areas more likely to reflect a patchwork of highly divergent genetic groups than widespread species found in regions that were affected by less extreme climatic and habitat change during the Pleistocene (but see Pfenninger and Schwenk 2007)?

Within widespread taxa that demonstrate marked genetic structure, addressing the factors limiting the spread of individual lineages is also of interest. For instance, in the case of the long-toed salamander, there is considerable variation in the distribution of individual lineages, raising questions as to why some lineages were able to colonize a much bigger area than others (*e.g*. the North-Central lineage, Figure 2.5) and whether these lineages might have been able to eventually occupy the entire contemporary range of the long-toed salamander if the other lineages were not present. Data on dispersal barriers, ecological

divergence and hybridization dynamics between the individual lineages are necessary to address these questions.

Finally I note that not all widespread taxa demonstrate strong phylogeographic structure. In the case of amphibians, some of the most-widespread species demonstrate remarkably little genetic differentiation across their range (*e.g*. *Ambystoma maculatum*: Zamudio and Savage 2003*; Lithobates [Rana] sylvatica*: Lee-Yaw *et al*. 2008; *Gastrophryne carolinensis*: Makowsky *et al*. 2009), including one of the few others to be widely distributed throughout the Pacific Northwest (*Anaxyrus [Bufo] boreas*: Goebel *et al.* 2009). Continued advances in genomic techniques will not only allow us to better survey taxa and thus identify true outliers with respect to range size, but will provide new means for testing ecological and evolutionary hypotheses for the remarkable variation observed in species' distributions.

2.6 Acknowledgements

For help with sampling and discussion, I thank Moses Michelsohn, Meghan Cooling, Nina Lobo, Stephen DeLisle, Matt Robinson, Talia Sechley, Leslie Anthony, Bill Leonard, Tom Titus, David Pilliod, Jay Bowerman, Ben Crabtree, Jason Weir, Bruce Maxwell, Lisa Hallock, Wes Savage, Caren Goldberg, Jason Irwin, Andy Giordano, Andrew Storfer and Kim Pearson. Mark Thompson provided several critical samples. Additional samples were provided by the Museum of Vertebrate Zoology (University of California, Berkeley). I appreciate the support of Parks Canada and in particular that of Lisa Larson, Barb Johnson, Charlie Pacas and Ward Hughson. For help with AFLP data collection, analysis and discussion, I thank Alan Brelsford, Nils Arrigo, Andrew Clarke, Matt Siegle, David Toews, Julie Allen, Rich FitzJohn and Kevin Omland.

Table 2.1. AFLP primer combinations used in survey of phylogeographic structure in the long-toed salamander. The first primer combination listed was that used to generate AFLP profiles for all individuals; the remaining five combinations were used to generate additional data for a subset of 70 individuals.

EcoRI primer (*NNN-3')	MseI primer (†NNN-3')	Dye	Number of Polymorphic Fragments (after filtering)
AGCA	CATA	NED	177
AGCG	CATC	PET	178
AGCA	CATG	NED-2010	95
AGCG	CATA	PET-2010	105
AGCT	CATA	VIC-2010	101
AGCC	CATC	FAM-2010	95

EcoRI primer: GACTGCGTACCAATTC*

MseI primer: GATGAGTCCTGAGTAA†

Table 2.2. Average pairwise AFLP distances (number of bands that differ in presence/absence) within (diagonal) and between groups.

	Coastal-Cascade	Central-Oregon Highlands	Rocky Mountains	North-Central
Coastal-Cascade	117.0			
Central-Oregon Highlands	133.3	110.0		
Rocky Mountains	128.3	126.8	102.1	
North-Central	128.0	124.7	106.3	90.3

Table 2.3. Average uncorrected (top) and TN93-corrected (Tamura and Nei 1993; bottom) pairwise sequence divergence long-toed salamander lineages identified using statistical parsimony. Diagonals show average uncorrected/corrected values within lineages.

	Santa Cruz	Southwestern Oregon	Coastal-Cascade	Central Oregon Highlands	Rocky Mountains	North-Central
Santa Cruz	NA/NA*	0.030	0.034	0.052	0.049	0.043
Southwestern Oregon	0.032	0.018/0.019	0.024	0.043	0.047	0.037
Coastal-Cascade	0.037	0.025	0.009/0.009	0.038	0.042	0.032
Central Oregon Highlands	0.058	0.048	0.042	0.006/0.006	0.049	0.029
Rocky Mountains	0.056	0.056	0.046	0.056	0.009/0.010	0.038
North-Central	0.048	0.048	0.035	0.031	0.041	0.012/0.012

*The putative Santa Cruz subspecies was represented by a single individual in the analysis and is shown in pairwise comparisons for illustration purposes only.

Table 2.4. Overlap between long-toed salamander lineages and putative glacial refugia in the Pacific Northwest (see Figure 2.1 inset for map).

Lineage	Putative Refugia
Coastal-Cascade	Olympic Peninsula (Soltis *et al.* 1997) Columbia River Drainage (Wagner *et al*. 2005; Steele and Storfer 2006) Coastal Mountains (Godbout *et al.* 2008) Klamath-Siskiyou Mountains (Soltis *et al.* 1997; Kuchta and Tan 2005; Steele and Storfer (2006)
North-Central (Northern and Southeastern)*	Clearwater River Drainage (Carstens *et al.* 2004) Edge of Columbia Plateau (Godbout *et al.* 2008)
Central Oregon Highlands	Blue-Wallowa Mountains west of the Snake River (see also Thompson and Russell 2005)
Rocky Mountain	Clearwater River Drainage (Carstens *et al*. 2004; Nielson *et al*. 2006) Salmon River Drainage (Carstens *et al.* 2005; Nielson *et al.* 2006)

*Note that the presence of divergent haplotypes of North-Central salamanders in an isolated population just east of the Cascades in Washington also points to long-term persistence of the lineage in the Columbia Plateau, an area that is largely a gap in the distribution of the species (and in the distributions of other species) apart from a few isolated populations.

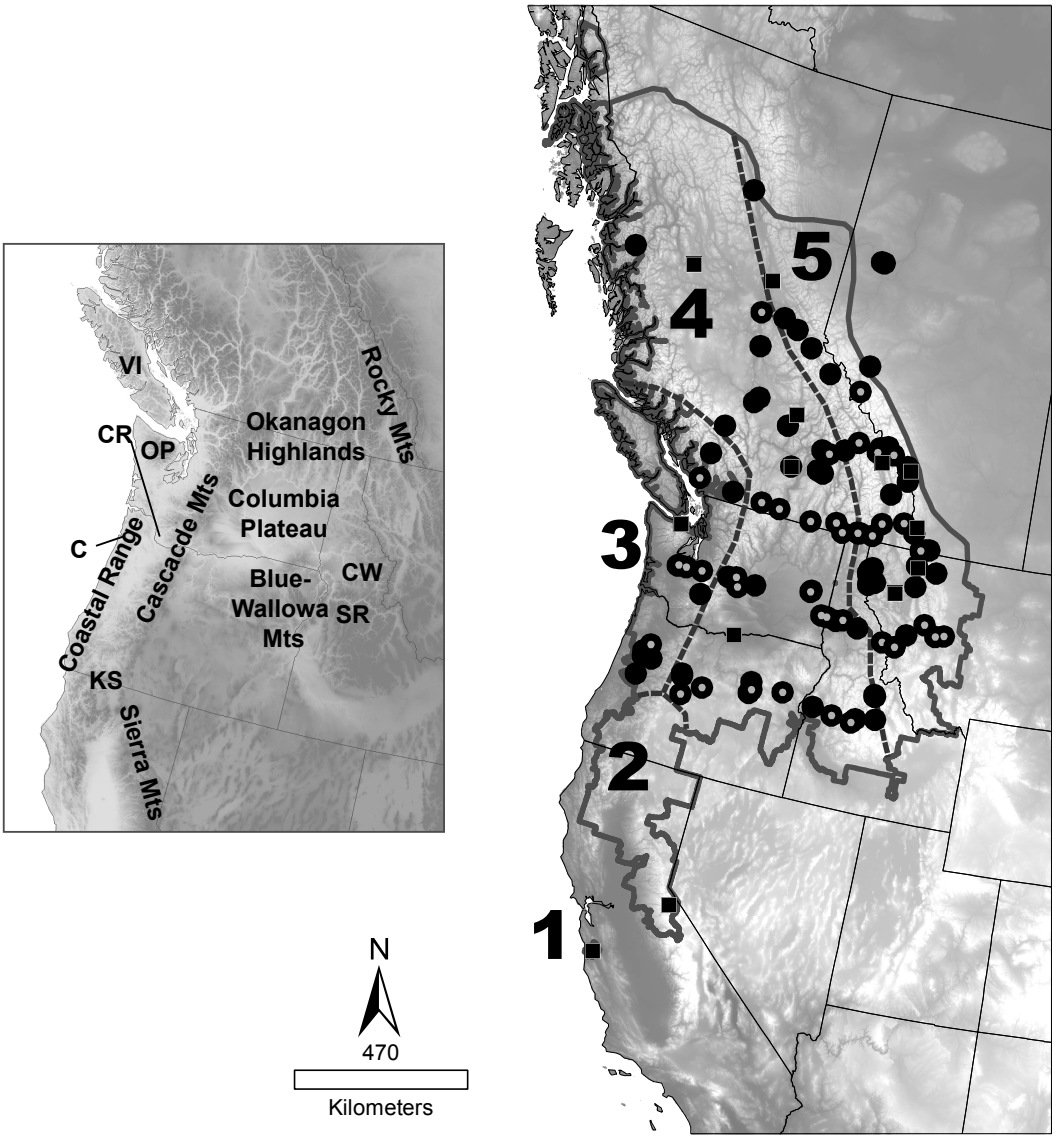

Figure 2.1. Genetic sampling in relation to current subspecies' boundaries (dashed lines): 1) *A. m. croceum* 2) *A. m. sigillatum* 3) *A. m. macrodactylum* 4) *A. m. columbianum* 5) *A. m. krausei*. The thick grey line shows the extent of the distribution of *Ambystoma macrodactylum*. Circles represent sites where both amplified fragment length polymorphism (AFLP) and mtDNA data were assayed (n = 1 to 11 individuals/site). Those sites with samples that were also included in a larger AFLP data set (see Methods; Figure 2.3) are indicated with a small grey circle within the larger black circle. Squares represent sites where only mitochondrial DNA (mtDNA) sequence data were collected (n = 1 to 4 individuals/site). Inset: Major geographic features and the locations of previously-reported glacial refugia as follows: VI = Vancouver Island, OP = Olympic Peninsula, CR = Columbia River Drainage, C = Coastal Mountains, CW = Clearwater River Drainage, SR = Salmon River Drainage, KS = Klamath-Siskiyou Mountains.

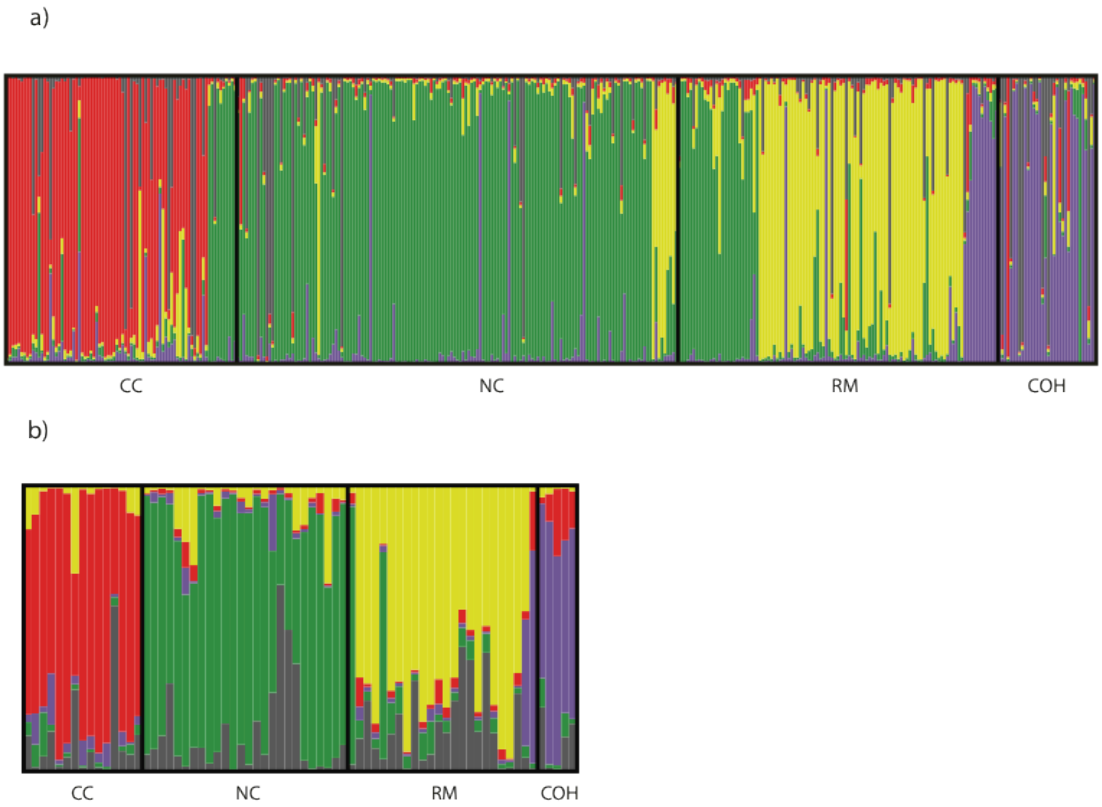

Figure 2.2. Results from STRUCTURE analyses of AFLP data: a) 177 AFLP loci for 378 long-toed salamanders and b) 751 AFLP loci for a subset of 70 individuals. Individuals are represented as bars (see Table A1 for population order for a), with the posterior probability of assignment to each of 5 clusters (the optimal value of K for both datasets) represented by the different colours. Mitochondrial (mtDNA) group membership, based on the haplotype network analysis (see Figure 2.4) is indicated below the plot with mtDNA groups separated by thick black lines. CC=Coastal-Cascade, NC=North-Central, RM=Rocky Mountains, COH=Central Oregon Highlands. Individuals are lined up such that those closest to the breaks separating mtDNA groups come from populations closest to the contact zones between lineages (Appendix A1).

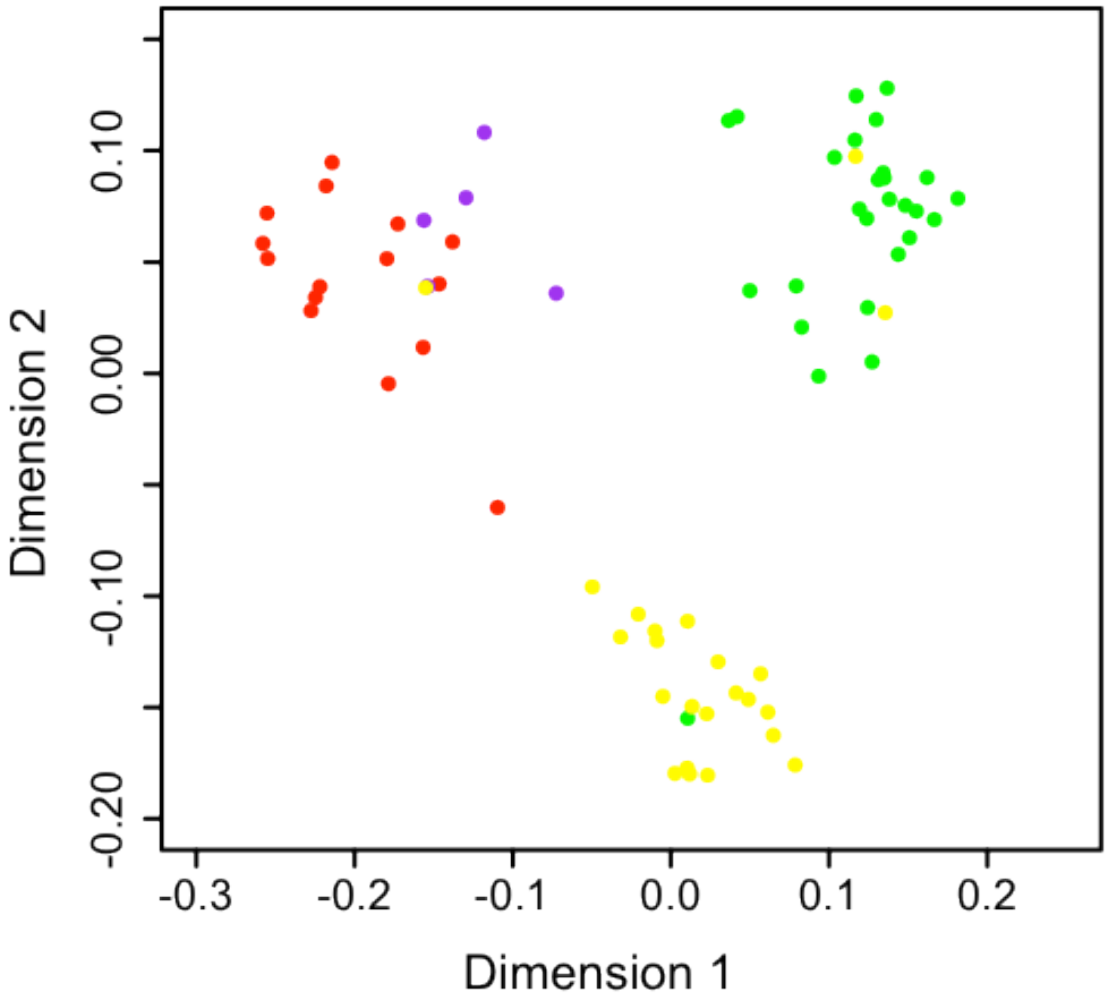

Figure 2.3. Multidimensional scaling plot of AFLP variation among 70 long-toed salamanders based on Jaccard distances calculated from 751 AFLP loci. Individuals are colour-coded according to mitochondrial DNA group membership from the haplotype network analysis (see Figure 2.4): North-Central (green), Rocky Mountain (yellow), Coastal-Cascade (red), Central Oregon Highlands (purple). All individuals demonstrating cyto-nuclear discordance were found along contact zones between lineages.

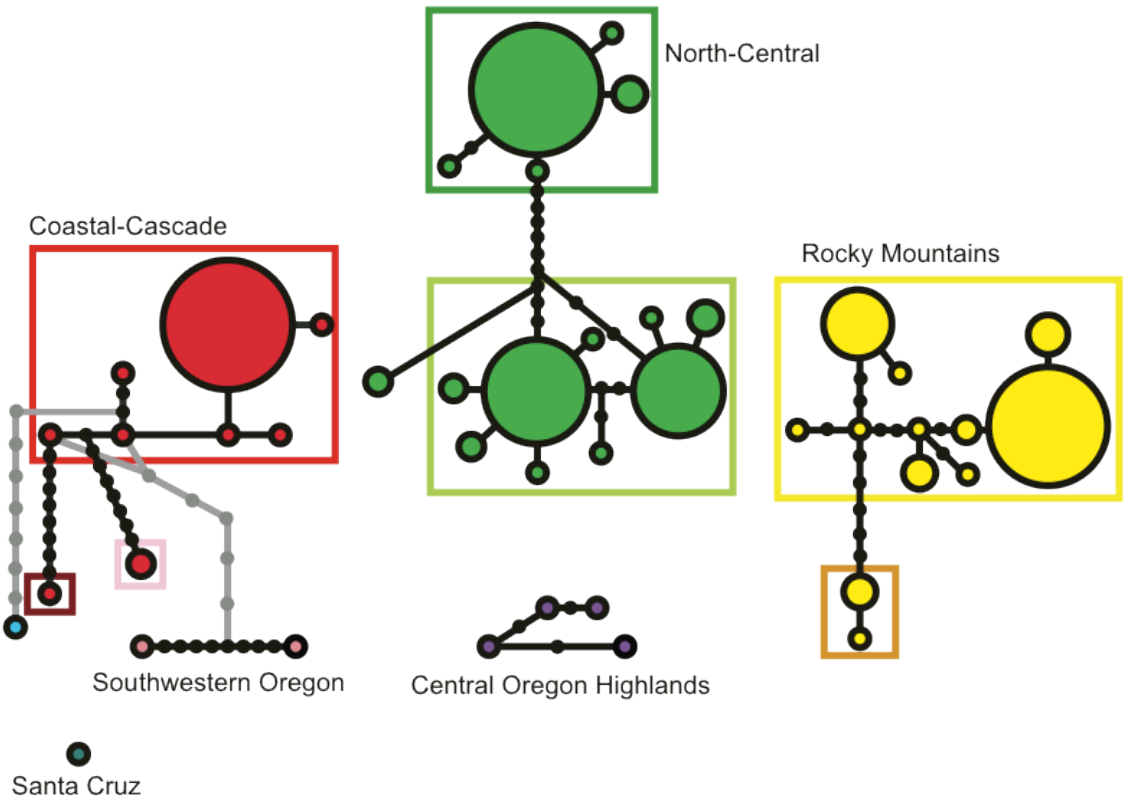

Figure 2.4. Haplotype networks based on 600 b.p. of mitochondrial *cytochrome b* from 142 Long-Toed salamanders. Network estimation was conducted in TCS (Clement *et al*. 2000) using statistical parsimony. Coloured circles represent haplotypes sampled from different parts of the species' range (see also Figure 2.5), with circle diameter proportional to frequency. Missing, intermediate haplotypes are shown as small black dots. Groups of haplotypes that are not joined by lines could not be connected within the 95% limits of statistical parsimony. Light grey lines show additional connections that are made when the analysis is repeated with 489 bp in order to include an individual from the Sierra Mountains in California (teal coloured haplotype). Coloured boxes around groups of haplotypes highlight divergent haplotypes groups within networks and show that further structure is detected within lineages (see also Figure 2.5).

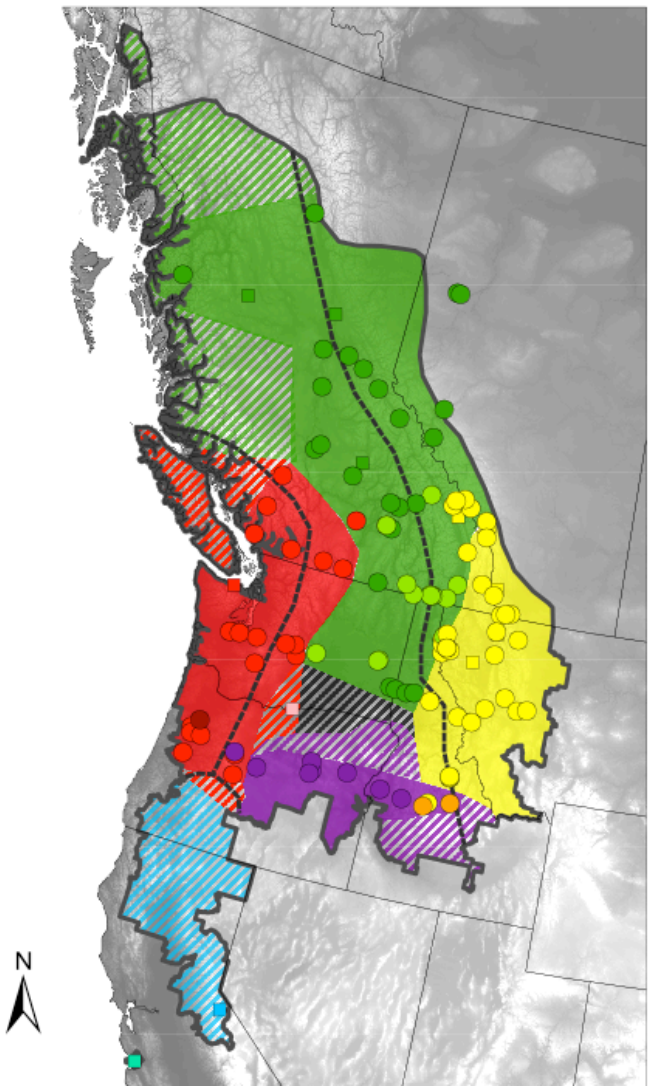

Figure 2.5. Approximate distributions of the major long-toed salamander lineages in relation to described subspecies' boundaries (dashed lines). Solid coloured areas denote the boundaries of the lineages based on the STRUCTURE analyses of the AFLP data (*i.e.* Figures 2.2 and 2.3) with the lineages coloured as follows: red = Coastal-Cascade, green = North-Central, yellow = Rocky Mountains, purple = Central Oregon Highlands. Striped colours represent regions where AFLP sampling was sparse and the distributions of the lineages are uncertain and based only on mtDNA, geophysiological features and/or previously described morphological variation (*e.g.* Ferguson 1961). Blue striping denotes the described range of the putative *A. m. sigillatum* subspecies and the teal point sample denotes the location of *A. m. croceum*. The black striped area represents an area where the species is very sparsely distributed. Sampling locations are coloured according the mtDNA group to which the majority of individuals from each site belong, with different shades denoting divergent groups within each network as indicated in Figure 2.4. Several instances of cyto-nuclear discordance are thus made evident and establish that hybridization occurs where the lineages come into contact.

Chapter 3: The influence of climate on parapatric range limits in the Long-Toed Salamander

3.1 Summary

Climate is thought to play a major role in limiting species' distributions; however, where range limits coincide with contact zones between closely related taxa, it can be hard to disentangle the relative importance of climate versus interspecific interactions in setting range limits. Here, I combine tests of niche overlap with ecological niche modeling to investigate the influence of climate on multiple, parapatric boundaries in the long-toed salamander (*Ambystoma macrodactylum*). I ask: Do niche models predict observed lineage boundaries, consistent with climate playing a role in setting these boundaries? Or, do model predictions suggest that each lineage could be more widespread, thus implicating other factors in generating these boundaries? I extend the existing framework for testing these hypotheses by evaluating variation in the climatic suitability of sites at the range limit and demonstrate an approach for incorporating multilocus genetic data into the delineation of the lineage boundaries that are tested. My results suggest that the different lineages of long-toed salamander are largely ecologically exchangeable and that suitable climatic space for each lineage exists beyond its present range. Some contact zones do coincide with localized areas of reduced climatic suitability. However, many sites at the range limit have suitability scores comparable to sites within the range of each lineage, making it unlikely that climatic barriers alone are driving range limits in the system.

3.2 Introduction

One of the most basic ecological properties of a species is its geographic distribution. Yet despite a broad literature pertaining to species' geographic range limits (reviewed by Gaston 2003), there are very few, if any, species for which a full explanation of the historical and contemporary factors that govern the range is available. The lack of such a treatise reflects, in part, the logistical difficulties of systematically evaluating the many ecological and evolutionary hypotheses for range limits (for review of hypotheses see Bridle and Vines

2007 and Sexton *et al.* 2009). Narrowing the list of alternative hypotheses for direct testing to those that are most pertinent for a given range is thus critical for any research program aimed at understanding species' range limits.

A classic dichotomy that can be made when characterizing range limits is between those that are primarily governed by abiotic factors and those that are largely mediated by the biotic environment (Gaston 2003). For a given species then, a useful first step is to establish whether observed limits reflect the geographic extent of the species' fundamental niche—that is, the limit of the species' physiological tolerance to abiotic conditions (*sensu* Hutchison 1957). Ideally this question is addressed through the use of transplant/enclosure experiments (Angert and Schemske 2005; Samis and Eckert 2009; Stanton-Geddes *et al.* 2012). However, such experiments are not feasible for many species and furthermore, may be difficult to conduct at appropriate spatial scales for understanding range limits.

Ecological niche modeling (ENM; Peterson *et al.* 2011) represents an alternative approach for assessing the extent to which species' distributions correspond with their abiotic niche limits. Ecological niche modeling involves using spatial data and geographic information systems (GIS) to extract information about the set of environmental conditions represented at locations where a species is present and generating a model that distinguishes these conditions from those found either where the species is absent or across a background dataset. Although the reliance on presence data to determine the environmental tolerances of a species means that niche models will almost certainly underestimate the fundamental niche (Peterson *et al.* 2011), when projected across geographic space, such models can still provide insight into areas that are suitable for a species based on conditions correlated with demonstrated population persistence. Comparing observed range limits to those predicted by the model thus speaks to the extent to which a species fills its potential niche with respect to the variables considered in the model. This approach has been widely used to address one of the most pressing questions in contemporary biogeography—the extent to which climate influences species' distributions (*e.g.* Graham *et al.* 2010; studies reported in Kharouba *et al.* 2013).

Parapatric range limits between closely related taxa represent a good opportunity to study the relative importance of climate in shaping distributional limits. Stable parapatric boundaries may correspond with sharp climatic breaks and simply reflect the edge of the

respective niches of the taxa involved (*e.g*. Endler 1977). However, such boundaries can also result from biotic interactions, the effects of which may or may not be influenced by climate. For instance, hybridization can result in the formation of parapatric range limits if hybrid individuals suffer reduced fitness causing hybrid zones to be demographic sinks (Barton and Hewitt 1985; Goldberg and Lande 2006). Reduced hybrid fitness may be mediated by external conditions, including adaptation to climatic conditions (reviewed by Keller and Seehausen 2012), but can also result from intrinsic genetic incompatibilities (*e.g*. Turner *et al*. 2012). Likewise, competition can act in concert with environmental gradients to create sharp parapatric boundaries (Case and Taper 2000; Price and Kirkpatrick 2009). However, the effects of competition can also be decoupled from climatic gradients if other factors influence the distribution of the resources for which taxa contend.

Ecological niche modeling provides a framework for assessing the importance of climate to parapatric range limits and thus for narrowing down these alternatives. Specifically, niche models can be built for both species involved in a parapatric boundary to assess the extent to which each fills its respective climatic niche. For a given species, if climate plays a role in limiting the range, the observed range limit is expected to correspond with either the predicted limit of suitable conditions (Figure 3.1a) or an area of low suitability (Figure 3.1b; see also Rissler and Apodaca 2007 and Glor and Warren 2008). If instead the niche model of a focal species suggests that suitable conditions extend continuously beyond the observed range limit (Figure 3.1c), factors other than climate—including interactions with the other species—may be more important drivers of range limits for the species.

I use this framework to ask whether climate plays a role in setting parapatric boundaries in a widely distributed species, the long-toed salamander (*Ambystoma macrodactylum*). The long-toed salamander is well suited for such a study for a couple of reasons. First, the species consists of multiple, distinct genetic lineages (Thompson and Russell 2005, Savage 2008; Lee-Yaw and Irwin 2012; Figure 3.2). Thus the system affords an opportunity to compare the results from different range limits in a closely related group of taxa. Similarly, the extensive latitudinal span of some of the parapatric boundaries in the species highlights an opportunity to explore variation in the relationship between range limits and climate (see also Werner *et al.* 2013). Apart from these considerations, the role of climate in shaping range limits in the system requires clarification. On the one hand,

populations demonstrate marked differences in thermal tolerances and life-history traits across elevations (Anderson 1967; Anderson 1972; Howard *et al*. 1983; Howard and Wallace 1985) and some authors have suggested that such variation may reflect differences between lineages (Anderson 1967). On the other hand many of the lineages occupy a similarly broad array of habitats and elevations within their respective ranges, raising questions as to whether there is any ecological specialization and divergence between lineages with respect to their climatic niches.

In this study, I use multilocus genetic data and a novel spatial approach to better delineate parapatric boundaries between the morphologically-similar long-toed salamander lineages. With the boundaries between the long-toed salamander lineages thus defined, I use GIS data and ecological niche models to ask: 1) Is there evidence to suggest that the different lineages of long-toed salamander have diverged with respect to their climatic niches? 2) Do parapatric boundaries reflect the limits of suitable climatic space or coincide with areas of reduced climatic suitability for the different lineages involved? 3) Is there variation in the relationship between range limits and climate along the length of a given parapatric boundary?

3.3 Methods

3.3.1 Delineating lineage boundaries

Accurate delineation of lineage boundaries is necessary to evaluate ecological divergence and the extent to which range limits coincide with the fundamental niches of the different lineages. I previously surveyed 378 individuals from 108 populations for 177 AFLP loci and used STRUCTURE (Pritchard *et al.* 2000; Falush *et al*. 2007) to estimate the fraction of each individual's genome derived from four main lineages (Coastal-Cascade, North-Central, Rocky Mountains and Central Oregon Highlands; Lee-Yaw and Irwin 2012; Chapter 2). In that study I provided a qualitative estimate of lineage boundaries. In this study, I use inverse distance weighting (IDW) spatial interpolation based on the STRUCTURE ancestry scores to provide a more quantitative assessment of lineage boundaries. Spatial interpolation allows for the prediction of the value of a continuous variable (in this case the "ancestry" of individuals within populations: see also Murphy *et al.* 2008) at unsampled locations based on existing observations. IDW specifically assigns values to unsampled

locations depending on values at sampled locations within a defined neighborhood size and weighted by distance. Of the many methods of interpolation available (reviewed by Li and Heap 2008), this method was reasonable for my dataset as it incorporates information from multiple populations (in contrast to nearest-neighbor methods) yet makes few assumptions about the data (*e.g.* ancestry scores were not normally distributed, thus violating the assumptions of more sophisticated geostatistical interpolation/ kriging methods).

I generated four interpolation surfaces—one for each of the genetic lineages. The interpolation procedure for a given lineage required every input location to have a single estimate of ancestry for the lineage in question. Prior to interpolation, it was thus necessary to summarize the STRUCTURE ancestry scores of all individuals for each lineage for each population. Simple averaging across individuals limits the ability to detect populations within contact zones. For instance, a sample of ten individuals, of which nine have 0% and one has 92% ancestry assigned to a focal lineage will have an average ancestry of 9.2% for that lineage, thus failing to account for the presence of an individual with a near-parental background in the population. Likewise, averaging across all individuals fails to distinguish populations comprised of admixed individuals from those comprised of parental-like individuals from two lineages. To overcome these limitations, I summarized population ancestry for each lineage as the average ancestry of those individuals with at least 50% ancestry assigned to the lineage in question. This cutoff amounts to assigning a value of zero (and thus "no presence" of the lineage) to populations where all individuals have less than 50% of their ancestry assigned to the lineage in question. Although arbitrary, this cutoff ensures two things. First, that at least one individual has a clear signature of ancestry for a given lineage before that population is counted within the range. Second, that no individual counts towards the ancestry estimates of more than two lineages—a scenario that is consistent with the parapatric nature of lineage boundaries in this system.

IDW Interpolation was performed using the Geostatistical Analyst tool in ArcGIS 10 with a variable neighborhood size of 3-10 neighbors and the default power function of two (corresponding to inverse distance squared weighted interpolation). Cell size was set to match the climate layers for the niche modeling (below). Following interpolation, I defined the range of each lineage as the region of space where population-level ancestry values are predicted to be at least 50%. The resulting boundaries were manually smoothed to remove

artifacts produced by noise in the AFLP dataset (*e.g*. holes in otherwise continuous range polygons and isolated polygons that are clearly inconsistent with the bulk of the data for each lineage). Further modifications were informed by the distribution of the monophyletic clades described by Savage (2008) and Lee-Yaw and Irwin (2012) based on nuclear and mitochondrial sequence data respectively (Figure B1). All modifications to the interpolated boundaries are described in Figures B1 and B2. Contact zones are naturally defined by this method as areas of overlap between the estimated ranges of the different lineages and are consistent with all existing genetic data.

I note that the AFLP dataset did not include populations in northern California or southern Oregon. Mitochondrial data (Savage 2008; Lee-Yaw and Irwin 2012) and two nuclear genes (Savage 2008) suggest that populations in these areas may represent a fifth lineage (Figure B1). Although other nuclear markers group these populations with populations to the north (Savage 2008; Figure A1), these populations have previously been described as a separate subspecies on the basis of morphology (Ferguson 1961). Thus, I treat populations in southwestern Oregon and California as a distinct lineage (hereafter referred to as *A. m. sigillatum* or *sig*). I delineated the boundaries of this lineage using a minimum convex polygon around all populations belonging to this putative lineage according to the data presented by Savage (2008). Populations in Santa Cruz (California) also represent a distinct lineage of long-toed salamander but are not considered here owing to their extremely limited distribution and geographic separation from the bulk of the species' range.

3.3.1.1 Locality data

Long-toed salamander locality records were compiled from other amphibian researchers, published papers, government databases and museum records. Only records with latitude/longitude data were included in the final locality database. Several records from each source were manually verified in Google Earth to ensure accuracy and to verify correct assignment of the geographic coordinate system and projection of the data. The final locality dataset consisted of 4821 records of the species across its range. I assigned these records to lineage using the boundaries derived above.

3.3.1.2 Climatic space occupied by different Long-toed Salamander lineages

To determine whether the different lineages of long-toed salamander occupy unique environmental space, I conducted a principal components analysis (PCA) based on 25 of the climatic variables (Table 3.1) available from ClimateWNA (Wang *et al.* 2012). These variables describe annual and seasonal conditions for western North America for the period between 1961 and 1990 at a 1-km^2 resolution. To further explore the extent of niche divergence between pairs of parapatric lineages, I used the multivariate test of niche overlap introduced by McCormack *et al*. (2010). Specifically, for each pair of parapatric lineages, I extracted climate data from all known salamander localities (including the contact zone between the two lineages) and from 1000 random (*i.e.* background) points from the allopatric range (*e.g*. that part of range in Fig. 3.2 that did not overlap with other lineages) of each of the two lineages as defined above. I conducted a PCA using these data and extracted the first three PC axes. The mean difference between allopatric sites of the two lineages along each of these axes was then calculated and compared to a null distribution of differences based on the background data along the same axis. The null distribution was generated by calculating the mean difference between 100 background points from within the range of each of the two lineages for 1000 bootstrap replicates of their respective background points. Lineages are considered more or less divergent along a given dimension of niche space than expected based on geography alone if the observed difference between lineages falls in the tails of the null distribution. PCAs and the tests of niche overlap were conducted in R (*stats* package: Becker *et al*. 1988).

3.3.1.3 Climate data for niche modeling

For purposes of building niche models, I focused on twelve climatic variables from the ClimateWNA dataset that are highly relevant to temperate amphibians. Specifically, I chose variables that are expected to impact hibernation and spring emergence (mean coldest month temperature, mean winter minimum temperature, precipitation as snow, mean winter precipitation, number of frost-free days), migration and breeding (spring precipitation, mean spring temperature), larval development (mean spring and mean summer temperature) and survival (mean annual precipitation, summer precipitation; maximum and mean summer temperature; summer heat to moisture index).

I extracted values for all climate variables from 1000 evenly spaced points across the range of the long-toed salamander. Using the *cor* function in R (*stats* package: Becker *et al*. 1988), I then calculated Pearson correlation coefficients between pairs of variables. For highly correlated variables (r>0.85), I retained only the variable that I deemed would have a more direct impact on population persistence (mean annual precipitation over mean spring [r=0.98] and winter precipitation [r=0.96]; minimum winter temperature over average temperature of the coldest month [r=0.98]; average summer temperature over maximum summer temperature [r=0.94] and average spring temperature [r=0.92]). All remaining pairwise correlation coefficients were <0.75, except for that between average minimum winter temperature and the number of frost-free days (r=0.79). I chose to retain both of these variables in the model as these variables are expected to influence different aspects of the species' biology (*e.g*. survival during hibernation versus duration of active period respectively). The final dataset thus consisted of seven climatic variables (listed as part of Table 3.1).

3.3.1.4 Generating and evaluating niche models

The choice of study extent when building niche models with presence-only data is of critical importance (VanDerWal *et al.* 2009; Anderson and Raza 2010; Barve *et al.* 2011). In particular, presence-only modeling algorithms use information from the background range of a species to differentiate the specific conditions that underlie occurrence records (Phillips 2008). If the study extent is set such that background points are drawn from a much larger region than that encompassing the occurrence records of the species, the resulting model may overfit the locality data, reflecting coarse-grain environmental variation that differentiates the general region of occurrence to the region of non-occurrence (Phillips 2008). More generally, background data or pseudo-absences for calibrating niche models should not be drawn from areas where the species is absent due to potential dispersal limitations or biotic interactions (*i.e.* areas where the species of interest is out of equilibrium with respect to climate: Peterson *et al*. 2011). Thus I restricted the extent of background sampling to the range of the lineage being modeled and then applied the model across space to generate the final predictions for that lineage (Phillips 2008; Anderson and Raza 2010).

In addition, because the niche models are being used to predict habitat suitability within the contact zones between lineages, I avoided *a priori* assumptions about the suitability of contact zone populations by excluding the contact zones from model calibration. I note that the decision to exclude contact zones may influence model predictions if the long-toed salamander's range as a whole includes environmental conditions that are found within the contact zones but not within the allopatric boundaries of the lineage being modeled. I explore this possibility when evaluating model transferability (see MESS analysis below) and assess the consequences of this decision by rerunning the analysis with contact zones included.

The maximum entropy algorithm employed by MaxEnt (Phillips *et al.* 2006) was used to construct ENMs for each lineage. I used an Albers Equal Area projection for all input environmental layers (Elith *et al.* 2011). For all models, duplicate locality records were removed prior to model generation. Models were constructed using hinge features alone (Phillips and Dudik 2008; Elith *et al.* 2011) with automatic setting of the regularization parameter (*i.e*. 0.5 based on Phillips and Dudik 2008). Variable importance was assessed using jack-knifing. Relatively few locality records were available from remote northern regions of the species' range. This data limitation generated latitudinal differences in sampling effort across the range of the North-Central lineage (Fig 2). To correct for this sampling bias when calibrating the model for this lineage, I provided MaxEnt with a bias grid (Elith *et al*. 2010) reflecting the probability of sampling based on distance to nearest population centre.

Prior to using a model to make predictions, it is necessary to evaluate model performance and accuracy. A standard approach is to divide locality data into a training set and a testing set. The training set is used to generate the ENM and the testing set, having been withheld from model parameterization, is used to assess model performance based on metrics such as area under the curve (AUC of receiver operator plots). Phillips *et al.* (2006) have noted the potential for model evaluation conducted in this way to be sensitive to the random split of the locality data. To overcome this limitation, it is common for investigators to employ *k*-fold cross validation, splitting the locality data into *k* folds and successively calibrating and evaluating the model with different folds held out for this evaluation step. This approach is useful for exploring potential variance in model performance arising from

the specific assignment of locality data into training versus testing sets. However, one question that is not addressed by this approach is how well models calibrated using the presence data outperform models based on random locations from throughout the study area; that is, the extent to which the locality data are actually informative. The random null models procedure suggested by Raes and ter Steege (2007) offers a way to address this question.

Here I combine these evaluation approaches. I use AUC as the test statistic. AUC assesses the ability of a model to discriminate between presence data and background data (relative to random classification of data). I use 5-fold cross validation to assess variation in the ability of models based on different partitions of the locality data to discriminate between withheld presence data and background data. For each fold, I additionally compare AUC values to a null distribution of values generated from 99 random locality datasets of the same size. The AUC value of a given model (and thus the extent to which that model outperforms a random model) is considered statistically significant if it falls outside the 95th quantile of this distribution. Only models with an AUC score ≥0.75 (see Swets 1988) and that passed this test were used to generate predictions across space. MaxEnt and the tests described here were implemented in R using the *dismo* (Hijmans *et al*. 2012) and *raster* (Hijmans and van Etten 2012) packages, as well as base functionality.

3.3.1.5 Model predictions across space

My goal was to use niche models calibrated from each lineage to make predictions about the climatic suitability of areas occupied by adjacent lineages. Model transfer across space (or time) in this way requires some care as extrapolation to conditions beyond those used to calibrate the model may be inappropriate (Phillips *et al.* 2011). To avoid extrapolation, I generated multivariate environmental similarity surfaces (MESS) for each lineage following the method outlined by Elith *et al*. (2010). These raster surfaces reveal areas beyond the boundaries of each lineage where values for one or more of the environmental variables fall outside of the range of values of the locality and background data used to calibrate the model. In this way, I was able to identify cells that would require extrapolation and exclude them from the final prediction map for each lineage. After identifying the appropriate extent of model transfer, I generated prediction surfaces for those

models from the k-fold cross validation procedure that passed the evaluation tests and calculated an average prediction surface for each lineage.

3.3.1.6 Climatic suitability across range boundaries

To determine whether range limits correspond with areas of reduced climatic suitability, I conducted an analysis of variance (ANOVA) for each lineage comparing the relative suitability of sites within the range to the suitability of sites within contact zones and beyond the range (*e.g*. Figure 3.1). Thus for each parapatric border (consisting of allopatric populations of both species and a contact zone between them), I conducted two analyses, first comparing the suitability of populations on either side of and within the contact zone based on the niche model of one lineage and then based on the niche model of the other. For comparisons involving the most extensive parapatric boundaries, I included a term describing the interaction between lineage of origin (Allopatric focal lineage, contact zone or Allopatric sister lineage) and latitudinal region (North, Mid and South) in the ANOVAs to determine whether the association between the range limit and climatic suitability varies across space. Localities were assigned to latitudinal region by dividing the maximum combined latitudinal span of both northern contact zones into three equal parts. As the data violated the assumptions of ANOVA in several cases, I used permutation tests to obtain p-values using the *lmp* package (Wheeler 2010) in R. I focused solely on parapatric boundaries running north-south in this study as the boundaries between the western and *A. m. sigillatum* lineages and between the North Central and Central Oregon Highlands lineages are more poorly defined.

3.4 Results

3.4.1 Do the lineages occupy different regions of climatic space?

Results from the principle component analysis show some differences in the climatic space occupied by the different lineages of long-toed salamander (Figure 3.3). The first three components had eigenvalues greater than one and explained 65.3%, 22.8% and 5.4% of the variation respectively (Table B1). PC1 was positively correlated with most temperature variables and negatively correlated with the amount of precipitation as snow. The North-Central and Rocky Mountains lineages showed some but not complete separation along this

axis. PC2 was positively correlated with most measures of seasonal and annual precipitation and negatively correlated with temperature seasonality. This axis separated the two western lineages from more eastern lineages. Populations in the Central Oregon Highlands lineage overlapped completely with populations in the North-Central and Rocky Mountains lineages along these axes. However, this lineage showed some separation from these populations along PC3 (not shown), which is positively correlated with the amount of summer precipitation and negatively correlated with the summer heat to moisture index. Western populations from the north and those corresponding to *A. m. sigillatum* showed some separation along both PC2 and PC3 (not shown).

To further explore differences in the climatic space occupied by parapatric lineages, I ran additional PCAs with pairs of lineages and compared the amount of divergence between allopatric populations of each lineage to background environmental divergence. PC1 and PC2 explained 88.4 to 93.4 % of the total variance in these analyses (Table B1). Variables loadings along these axes were qualitatively similar to those observed in the PCA that included all lineages (Table B1). Results from the niche overlap test suggest that most pairs of parapatric lineages are generally either less divergent (*e.g.* niche conservatism) or no more divergent than expected based on differences between their background environments along both of these niche dimensions (Table B1).

3.4.2 Do range limits reflect the limits of suitable climatic space?

Niche models were characterized by high AUC scores (allopatric models: 0.76 to 0.91; models including contact zones: 0.76 to 0.89) with all replicates from the cross-validation procedure performing significantly better than models based on random points within the range of each lineage. Jackknife tests of variable importance suggest that climatic suitability is driven by different variables for the different lineages (Table 3.1)—although summer heat to moisture and mean minimum winter temperature were of high importance for several models. With the exception of the models for the Central Oregon Highlands lineage, similar variables were important to the niche model of any given lineage regardless of whether the model was based on allopatric populations or incorporated populations from the contact zones (Table 3.1).

Models were projected across space to determine whether range limits reflect the limits of the climatic niche for the different lineages (*e.g*. Figure 3.1a). Maps of model predictions indicate that suitable climatic conditions for each lineage occur well beyond their existing range limits (Figure 3.4). This pattern was observed regardless of whether models were based only on allopatric populations or incorporated populations from contact zones (*e.g*. compare Figure 3.4 to Figure B3). Incorporation of contact zones lead to a general increase in the relative suitability scores across the study area for most lineages. However, patterns of relative suitability across space were qualitatively similar between the two types of models and I present remaining results for the allopatric models, highlighting only those places where results differ.

3.4.3 Do range limits coincide with areas of low climatic suitability?

I asked whether range boundaries—namely the contact zones between lineages—correspond with areas of low climatic suitability (*i.e.* Figure 3.1b). Lineage of origin (allopatric focal lineage, contact zone or allopatric sister lineage) was significant in all but two of the ANOVAs comparing average suitability among populations (Table 3.2). Examination of 95% confidence intervals suggests that declines in average suitability within contact zones contributed to this result in several cases (Figure 3.5). For instance, when niche models are calibrated using allopatric populations, average suitability is consistently lower within contact zones relative to allopatric populations at comparable latitudes for the North-Central and Coastal-Cascade lineages. As expected, when populations from contact zones are used to inform the models during calibration, suitability in the contact zones increases. Nevertheless, a pattern of decreased suitability at range limits relative to conditions within the range is maintained in several of the comparisons involving these lineages (Figure B4).

In some comparisons, suitability within the range of the sister lineage is also low relative to conditions within the range of the focal lineage (Figure 3.5). For instance, much of the range of the North-Central lineage appears to be of reduced climatic suitability for the Coastal-Cascade lineage, especially when models are based on the allopatric range (Figure 3.4). Likewise, regions with suitable climatic conditions for *A. m. sigillatum* are separated from the observed range of this lineage by areas of relatively low suitability that extend over much of the range of the Central Oregon Highlands lineage (Figure 3.4).

In other cases, range limits are clearly not associated with declines in climatic suitability. For instance, average suitability for the Central Oregon Highlands increases within its parapatric boundary with the Rocky Mountain lineage (Figures 3.5 and B4). Likewise, for the Rocky Mountain lineage, contact zones (and often the region beyond the range limit) coincided with equally (or slightly more) suitable conditions with respect to those found within the range (Figure 3.5); although this pattern is not observed in southern regions when the broad contact zone with the Central Oregon Highlands lineage is included in the data used to calibrate the niche model (Figure B4).

3.4.4 Does the association between climate and range limits vary in space?

I took advantage of the extensive parapatric boundaries between the North-Central lineage and the Costal-Cascade and Rocky Mountains lineages to ask whether the relationship between climatic suitability and range limits is consistent across different latitudes. The interaction term between lineage of origin and latitudinal region was significant in all relevant ANOVAs suggesting that there are differences in the nature of the relationship between population of origin and suitability among latitudinal regions (Table 3.2). However, examination of Figure 3.5 suggests that this result largely reflects latitudinal variation in the extent to which populations occupied by sister lineages are suitable for the focal lineage as contact zones demonstrated striking similarities in relative suitability across latitudes for each lineages. As a final point, I note that there is a substantial amount of variation around the mean suitability score within all groups at all latitudes. Thus, despite changes in mean suitability across range limits, there are locations within and beyond the contact zones of each lineage that are equally or more suitable for the lineage than locations within the range.

3.5 Discussion

I set out to determine whether climate plays a role in shaping parapatric range limits in one of the most widely distributed amphibians in the Pacific Northwest. Surprisingly, most of the different lineages of long-toed salamander demonstrate more similarity with respect to their climatic niches than is expected based on the geographic regions that they occupy. Likewise, suitable climatic space exists beyond the range of each lineage, refuting the first hypothesis in Figure 3.1. Consistent with the second hypothesis presented in Figure 3.1,

modest dips in climatic suitability corresponding with range limits suggest that local climatic barriers may impact some range limits in the system; however, the high suitability of some sites within contact zones suggests that these barriers are not absolute. Thus other barriers acting alone or in concert with these regions of low climatic suitability are likely reinforcing range limits in this system.

3.5.1 A role for climate in shaping parapatric range limits

My results suggest that the availability of suitable climatic space does not limit the distribution of individual long-toed salamander lineages in an absolute sense. Although the different lineages occupy somewhat different regions of climatic space (Figure 3.3), results from the niche overlap tests suggest that most pairs of parapatric lineages are either no more divergent along climate axes than their background environments or demonstrate niche conservatism (Table 3.1). Predictions based on ecological niche models corroborate these results, suggesting that extensive areas of suitable climatic space exist beyond range limits for all lineages (Figure 3.4). Thus the different lineages of long-toed salamander appear to be largely ecologically exchangeable and are not generally partitioned in space by major differences in their climatic niche.

Nonetheless, climate may still play a role in setting the boundaries between lineages. In many cases, range limits coincide with a reduction in average climatic suitability (Figure 3.5), suggesting that some sites within contact zones (and in some cases, well beyond the range limit) may be demographic sinks. That this pattern is observed repeatedly across latitudes for the North-Central and Coastal-Cascade lineages lends support to the role of climate in setting the range limits of these lineages. Furthermore, the greatest breakdown of lineage boundaries—that is, the most introgression—occurs in areas where the relative suitability of sites is high for one or both lineages. In particular, the broadest contact zones involve the Rocky Mountains lineage. In three of the four latitudinal regions examined, relative suitability across range boundaries is high for this lineage. Thus range boundaries are more easily blurred where climatic barriers are weak for one (or both, in the case of the boundary with the Central-Oregon Highlands lineage) of the lineages involved in a parapatric boundary. Finally, I note that the global range limits of *A. macrodactylum* are generally well

predicted by the niche models of the lineages most immediate to them (Figure 3.4) suggesting that climate may generally be important for range limits in this system.

That some range limits in the long-toed salamander coincide with areas of reduced climatic suitability is consistent with observations from studies of more narrowly distributed species. Rissler and Apodaca (2007) found that contact zones in *Aneides* salamanders in California coincide with areas of low suitability for some lineages, separating areas of high suitability from areas of intermediate suitability. Glor and Warren (2010) presented a method for testing whether contact zones coincide with ribbons of unsuitable habitat that are more extreme than what would be observed if range limits were randomly positioned on the landscape. Using this method, they demonstrated that a ribbon of particularly unsuitable habitat coincides with the boundary between *Anoles* lizards on Hispaniola (Glor and Warren 2010). Soto-Centeno *et al*. (2013) used this method to demonstrate that a contact zone between pocket gophers in the southeastern USA is also associated with a ribbon of unsuitable habitat. My results suggest that such climatic barriers may similarly impact the distribution of diversity in more widely distributed species.

3.5.2 A role for other range-limiting factors

Although some range limits coincide with areas of reduced climatic suitability, climatic barriers alone do not fully explain range limits in the long-toed salamander. The variance in suitability scores was large, with some sites in contact zones and beyond the range of each lineage occurring where climatic conditions are predicted to be highly suitable. Therefore range limits do not coincide with absolute gaps in suitable conditions. Furthermore, as already noted, climatic barriers do not readily explain range limits in the case of the Rocky Mountains lineage. Thus additional barriers are likely operating to maintain range limits in this system.

The range of the long-toed salamander, like other northern taxa, reflects recent colonization following the last glacial maximum (~15,000 ybp). It is possible that the system is simply out of equilibrium and that, given sufficient time, each lineage will expand to fill available suitable habitat producing large areas of overlap between lineages or hybrid swarms on the landscape. However, the elongated ranges of the different lineages demonstrate that sufficient time has passed for extensive range expansion in northern directions, suggesting

that the longitudinal ranges of the lineages are limited by factors other than time for colonization. In some instances, this pattern may reflect physical barriers to dispersal. For instance, the Columbia River clearly separates Coastal-Cascade and North-Central populations near Ellensburg in central Washington. The extent of introgression between the North-Central and Rocky Mountains lineages in southeastern British Columbia may similarly be influenced by this large river system (Lee-Yaw unpublished). However, physical barriers can be overcome in time and the different long-toed salamander lineages have clearly breached equally large dispersal barriers within their ranges. Furthermore, most contact zones are not associated with obvious dispersal barriers. Thus dispersal limitation is unlikely to be a major driver of range limits in this system.

Given the extent of niche overlap between lineages, interactions between lineages likely play a role in maintaining parapatric boundaries in this system. For instance, competition has been invoked in other salamanders to explain the exclusion of species from otherwise suitable habitat (Arif *et al.* 2007; Cunningham *et al.* 2009). To the best of my knowledge, there have been no studies examining competition between the different lineages of long-toed salamander. However, competition within ponds is known to have dramatic consequences for individuals in the system. For instance, competition amongst larvae has been reported to result in the production of large, cannibalistic morphs (Wildy *et al*. 2001). Although it is not known whether such morphs arise (or arise more frequently) in sympatric populations, the production of different larval feeding morphs has been observed in response to competition with conspecifics in another amphibian—spadefoot toads in the genus *Spea*. In that system, although there is quite a bit of plasticity, *S. bombifrons* preferentially develops into a larger cannibalistic morph whereas *S. multiplicata* develops into a smaller omnivorous morph in sympatry (Pfennig and Murphy 2000). This character displacement reduces competition between the species but comes at a fitness cost in terms of survival and fecundity for the smaller morph (Pfennig and Pfennig 2005). Such fitness costs arising from this or other forms of character displacement in response to competition at range boundaries in the long-toed salamander could conceivably reinforce range limits, especially if competition is more severe in areas of reduced climatic suitability.

The presence of admixed individuals on the landscape indicates that the lineages hybridize where they come into contact (Savage 2008; Lee-Yaw and Irwin 2012).

Hybridization can result in the formation of range boundaries if hybrid individuals suffer a fitness cost (Barton and Hewitt 1985; Goldberg and Lande 2006). Furthermore, range limits mediated by selection against hybrids are expected to settle in regions of low dispersal or population density (Barton and Hewitt 1985; Goldberg and Lande 2007). A leading alternative model for the maintenance of hybrid zones—the bounded hybrid superiority model, whereby hybrids demonstrate higher fitness than parentals in some environments (Moore 1977)—would also predict that range limits (*e.g.* hybrid zones) settle in places where the parental forms are not well adapted. Thus, regardless of the direction of selection, hybridization might be expected to lead to range limits that are attracted to localize areas of low climatic suitability. Both the effects of competition and hybridization on range limits in the long-toed salamander require further evaluation.

3.5.3 Niche conservatism and the maintenance of diversity

Although niche models and tests of niche divergence only speak to the variables examined and it is possible that long-toed salamander lineages differ with respect to other aspects of their ecology, my results suggest that the lineages are highly similar with respect to the most obvious dimensions of their climatic niche. Thus the long-toed salamander adds to a growing number of closely related taxa that demonstrate niche conservatism (Peterson *et al*. 1999; Peterson 2011). Apart from the implications for understanding geographic range limits, these examples stand in contrast to the current emphasis on divergent ecological selection, including climate adaptation (Keller and Seehausen 2012; Schnitzler *et al*. 2012), as the primary driver of speciation. For northern taxa in particular, genetic drift and/or mutation-order adaptation (*e.g*. reviewed by Schluter 2009) following the tracking of suitable conditions into allopatric refugia during the Pleistocene glaciations may have been just as important for divergence.

Of course the important question is whether these lineages (and their subsequent range limits) are likely to persist through time. Here climate may indirectly influence the ultimate fate of lineages. In particular, when taxa demonstrate niche conservatism, the distribution of suitable climatic conditions can influence the potential for gene flow between lineages (*e.g.* Arteaga *et al.* 2011). At one end of the spectrum, large regions of inhospitable conditions may separate areas with suitable climatic conditions, promoting allopatry and

facilitating further divergence (Wiens 2004; Kozak and Wiens 2006). Results for the long-toed salamander (as well as the *Anolis* lizards and pocket gophers mentioned above) demonstrate that narrowing the gap between areas of suitable climatic conditions can result in parapatric distributions among otherwise ecologically similar lineages. This proximity makes gene flow possible, especially when climatic barriers are incomplete as observed here. In turn, the consequences for diversification range from introgression across much of the genome (*e.g.* Wiens *et al.* 2006) to the accelerated development of reproductive isolation (*e.g*. through reinforcement: Servedio and Noor 2003; Pfennig and Pfennig 2009) to transgressive segregation (*e.g.* Chunco *et al*. 2012). Thus even when climate adaptation does not directly drive diversification, climate can have important consequences for the maintenance of diversity through its effects on species' geographic distributions.

3.6 Acknowledgements

I thank Tongli Wang for providing ClimateWNA data in raster format. Mark Thompson. Stephanie Chong was instrumental in organizing incoming emails and maintaining the locality database. I thank Mike Meitner for clarifying several GIS concepts and for help with ArcGIS software.

Table 3.1. Contributions of different climatic variables to niche models built in Maxent based on the allopatric (full) ranges of the different lineages of long-toed salamander.

Lineage		map	nffd	pas	ppt_sm	shm	tave_sm	tmin_wt
CC	Contribution (%)	11.86 (10.47)	1.83 (1.82)	10.38 (8.34)	3.64 (7.61)	48.25 (42.03)	5.51 (8.10)	18.52 (21.64)
	Permutation importance (%)	7.73 (10.96)	2.24 (4.28)	8.93 (7.41)	6.90 (9.04)	35.34 (27.21)	8.29 (13.05)	30.58 (28.05)
	Training gain if excluded	0.53 (0.46)	0.55 (0.48)	0.54 (0.47)	0.54 (0.46)	0.53 (0.47)	0.55 (0.47)	0.50 (0.43)
	Training gain if used in isolation	0.23 (0.17)	0.15 (0.13)	0.13 (0.10)	0.30 (0.23)	0.31 (0.22)	0.07 (0.07)	0.18 (0.17)
NC	Contribution (%)	2.79 (5.90)	1.68 (0.42)	2.58 (3.63)	0.87 (0.4)	43.24 (35.89)	35.76 (5.61)	13.07 (48.16)
	Permutation importance (%)	3.12 (6.14)	0.82 (4.32)	2.93 (12.47)	12.99 (2.59)	0.02 (6.82)	75.71 (42.35)	4.42 (25.3)
	Training gain if excluded	0.88 (0.67)	0.89 (0.67)	0.89 (0.67)	0.88 (0.68)	0.89 (0.68)	0.85 (0.64)	0.87 (0.63)
	Training gain if used in isolation	0.23 (0.11)	0.62 (0.34)	0.43 (0.15)	0.48 (0.34)	0.60 (0.38)	0.79 (0.48)	0.44 (0.43)
RM	Contribution (%)	34.05 (24.52)	5.05 (6.82)	15.01 (15.21)	1.58 (11.92)	2.59 (2.38)	7.75 (6.09)	33.97 (33.06)
	Permutation importance (%)	7.38 (15.35)	4.00 (5.65)	38.5 (27.86)	4.64 (11.12)	8.49 (5.37)	16.92 (14.58)	20.08 (20.07)
	Training gain if excluded	0.35 (0.25)	0.35 (0.26)	0.34 (0.26)	0.35 (0.26)	0.35 (0.27)	0.35 (0.26)	0.33 (0.22)
	Training gain if used in isolation	0.12 (0.09)	0.02 (0.02)	0.10 (0.08)	0.03 (0.06)	0.05 (0.06)	0.02 (0.03)	0.08 (0.07)
Sig	Contribution (%)	12.19	4.83	59.85	4.44	6.32	10.36	2.02
	Permutation importance (%)	5.13	14.91	55.04	5.3	0.53	15.74	3.33
	Training gain if excluded	1.09	1.09	1.02	1.08	1.10	1.00	1.08
	Training gain if used in isolation	0.28	0.35	0.65	0.21	0.41	0.40	0.34
COH	Contribution (%)	19.77 (16.25)	2.51 (5.78)	1.55 (25.54)	14.17 (0.11)	3.00 (44.17)	2.72 (4.67)	56.28 (3.48)
	Permutation importance (%)	23.91 (25.74)	5.23 (17.11)	2.02 (11.19)	27.64 (0.19)	10.69 (7.04)	2.28 (1.87)	28.21 (36.87)
	Training gain if excluded	0.47 (0.71)	0.52 (0.80)	0.53 (0.79)	0.49 (0.82)	0.53 (0.80)	0.52 (0.80)	0.45 (0.79)
	Training gain if used in isolation	0.12 (0.46)	0.20 (0.48)	0.09 (0.49)	0.12 (0.46)	0.08 (0.53)	0.07 (0.51)	0.29 (0.38)

map = mean annual precipitation; nffd = number of frost-free days; pas = precipitation as snow; ppt_sm = average summer precipitation; shm = summer heat to moisture index; tave_sm = average summer temperature; tmin_wt = average minimum winter temperature

Table 3.2. Effects of lineage and range position on average climatic suitability from niche models of long-toed salamander lineages based on their allopatric ranges*.

		Suitability based on niche model of western-most lineage in comparison			Suitability based on niche model of eastern-most lineage in comparison		
Boundary†	Source of Variation	SS	df	P	SS	df	P
CC – NC	Lineage of Origin	3.571	2	<0.001	1.353	2	<0.001
	Latitudinal Region	2.323	2	<0.001	0.144	2	0.345
	Lineage of Origin x Latitudinal Region	1.0972	4	<0.001	0.690	4	<0.001
NC – RM	Lineage of Origin	3.748	2	<0.001	0.295	2	<0.001
	Latitudinal Region	1.826	2	<0.001	5.302	2	<0.001
	Lineage of Origin x Latitudinal Region	0.555	4	<0.001	1.630	4	<0.001
sig – COH	Lineage of Origin	0.0184	2	0.75	1.64	2	<0.001
COH – RM	Lineage of Origin	1.680	2	<0.001	0.0581	2	0.5

*Significance evaluated using permutation tests

†CC = Coastal-Cascade lineage; NC = North-Central lineage; RM = Rocky Mountains lineage; sig = *A. m. sigillatum*

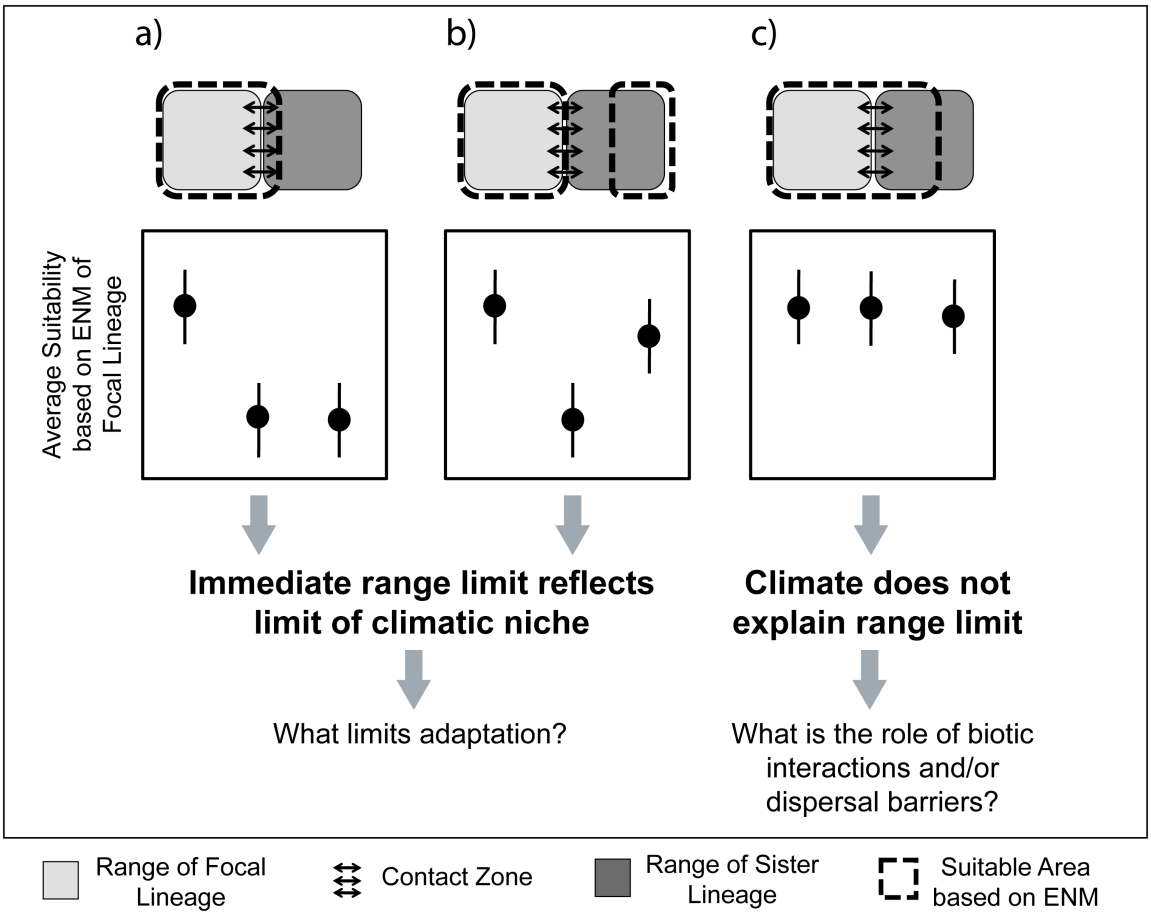

Figure 3.1. Framework for studying the importance of climate for parapatric boundaries using ecological niche models (see also Rissler and Apodaca 2007 and Glor and Warren 2010). A close match between observed range limits and the extent of suitable climatic space (a) or the extent of localized areas of suitable climatic space (b) predicted from ecological niche models is suggestive of a role of climate in setting the range limits of a focal lineage. Average suitability is expected to decline for populations at and/or beyond the range limit in these cases. Alternatively, range limits that fall short of the extent of continuous suitable climatic space, with no marked decline in the average suitability of populations across range boundaries (c), are likely maintained by factors other than climate. Applying this framework to both taxa involved in a parapatric boundary sheds light on any asymmetry in the role of climate on range limits. Alternative outcomes can be used to guide future research efforts to fully understand the mechanisms maintaining the parapatric boundary.

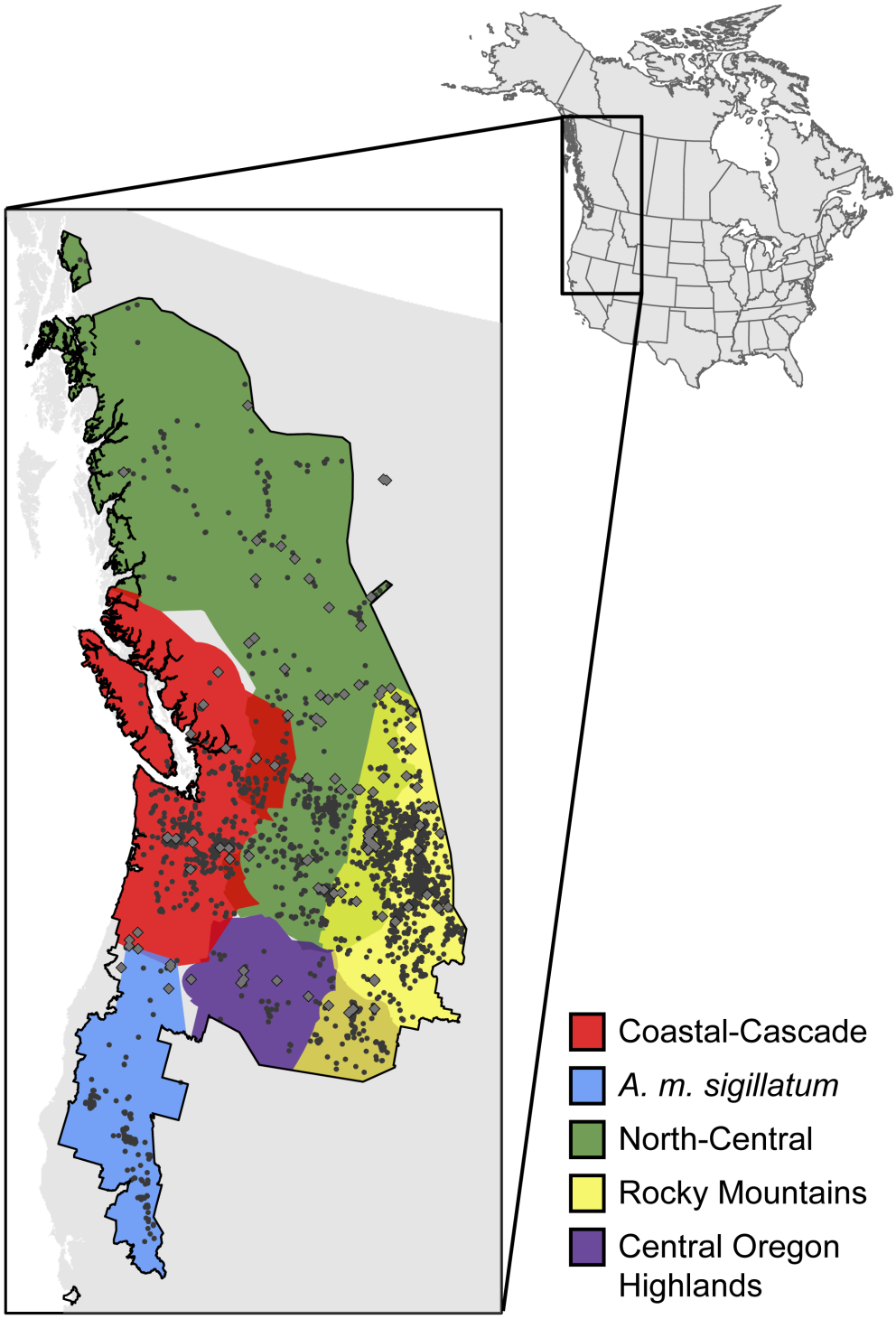

Figure 3.2. Distribution of the long-toed salamander in western North America and locality data used to build ecological niche models. Except for A. m. sigillatum, lineage boundaries are based on inverse distance weighting spatial interpolation of the STRUCTURE ancestry values observed at the subset of locations shown as grey diamonds. The range of A. m. sigillatum lineage is based on sequence data from Savage (2008).

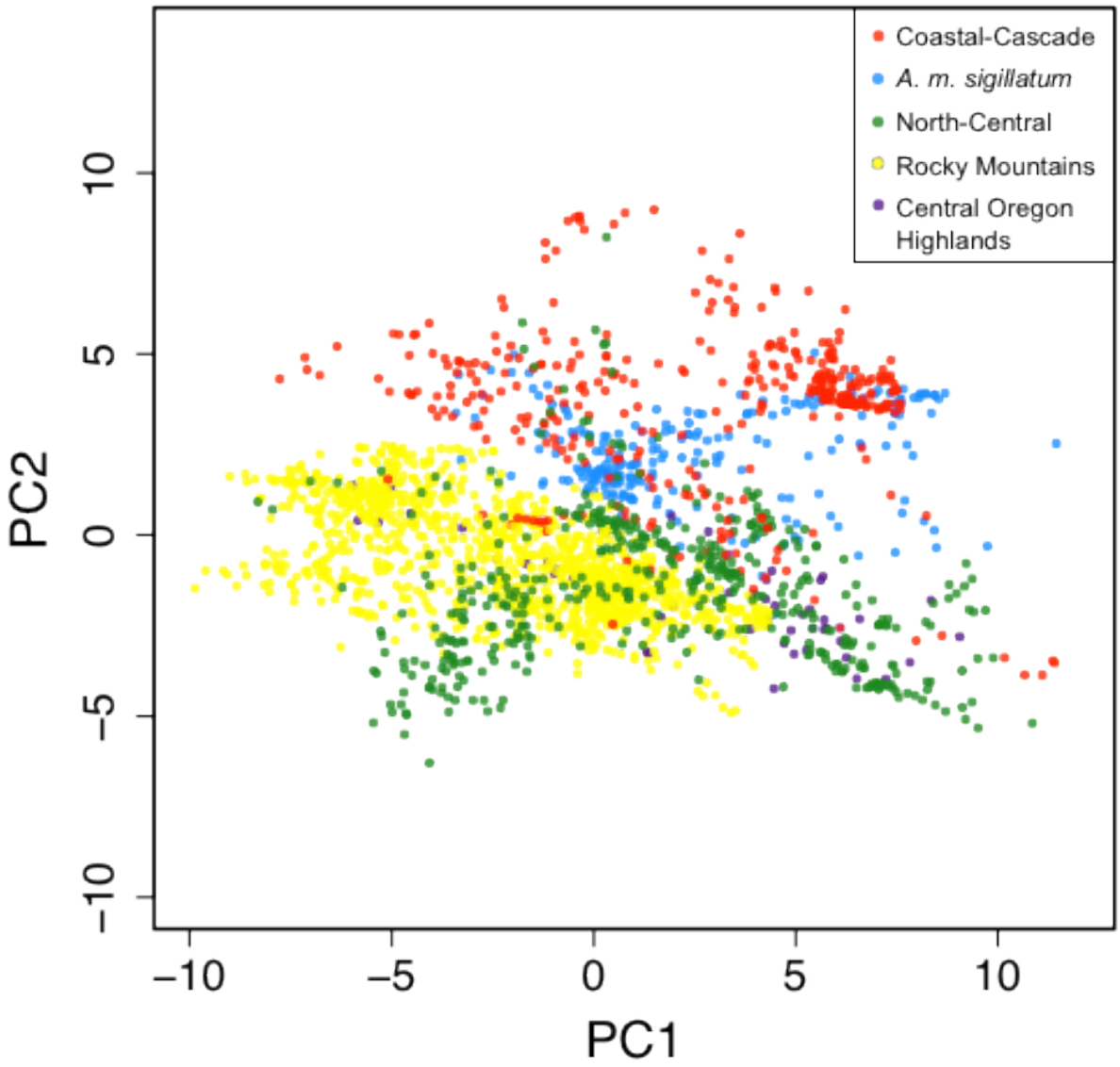

Figure 3.3. PCA of climatic variation among long-toed salamander localities based on 25 climatic variables. Different lineages are denoted by different colours. Note that the PCA included populations from the contact zones but only allopatric populations from the range of each lineage are plotted here for clarity.

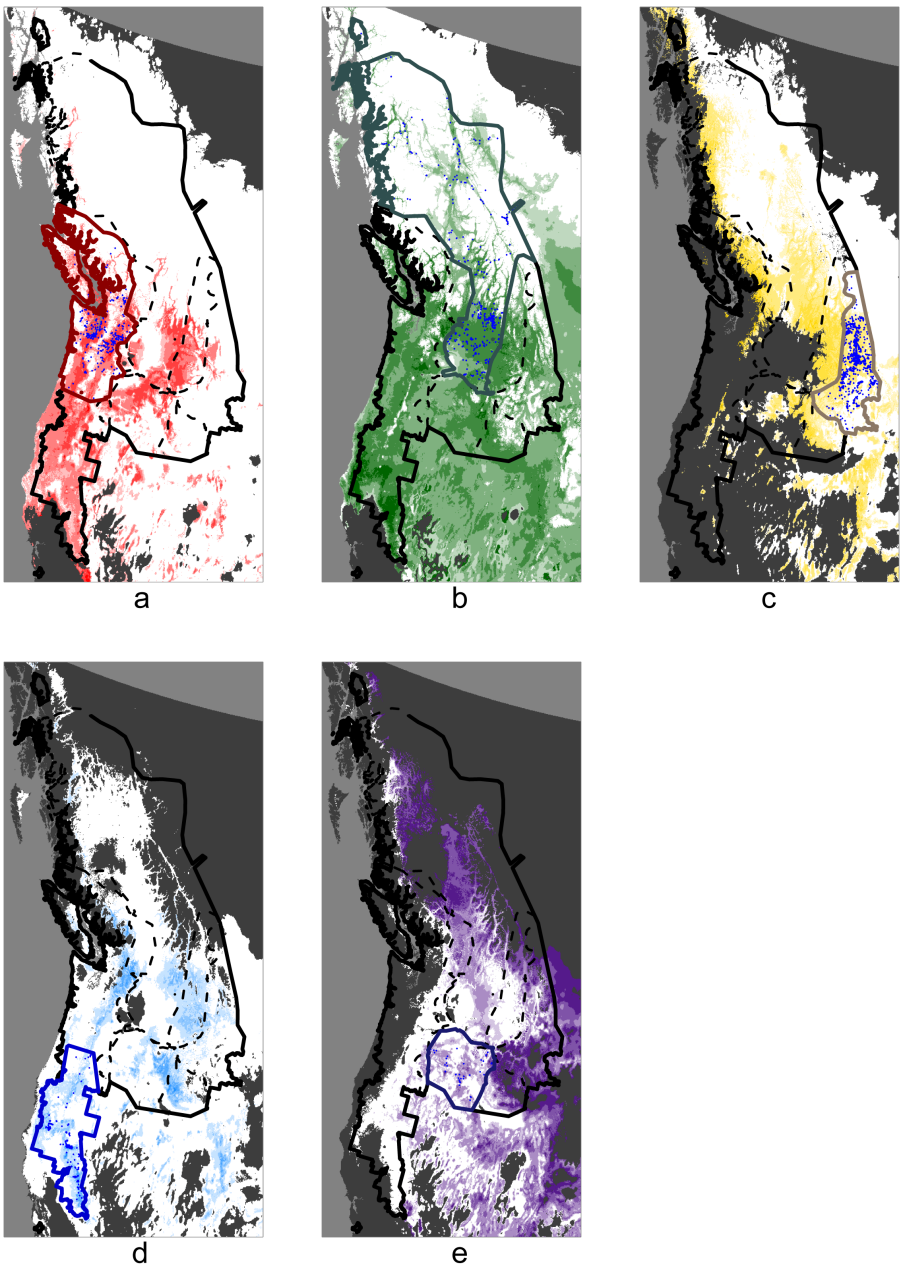

Figure 3.4. Suitability maps from ecological niche models based on the allopatric boundaries of the different lineages of long-toed salamander: a) Coastal-Cascade b) North-Central c) Rocky Mountains d) *A. m. sigillatum* e) Central-Oregon Highlands. Darker colours indicate higher suitability. All areas where suitability was lower than 95% of the sites within the focal lineage's range were assigned a value of zero (white) for plotting purposes. Small blue dots represent sampling locations. Dark grey areas correspond to places where the range of one or more climatic variables was outside that used to calibrate the model and thus where model extrapolation would have been necessary to make predictions (*e.g.* results from the MESS analysis). The solid black line delimits the boundaries of *A. macrodactylum* with dashed lines showing the lineage boundaries as per Figure 3.2. Solid coloured lines delineate the allopatric range of the focal lineage.

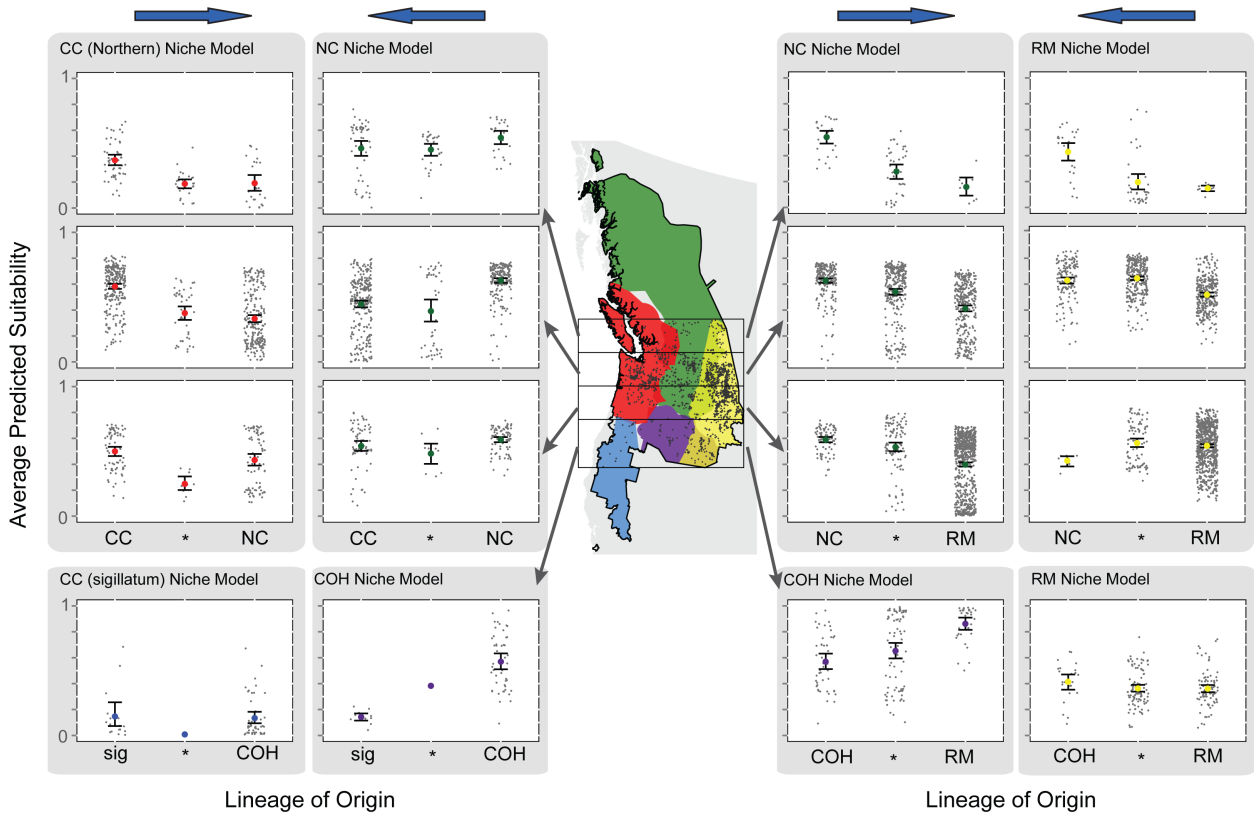

Figure 3.5. Average climatic suitability (with 95% CI based on 1000 bootstrap replicates) of populations (shown as grey dots in plots) across range boundaries at different latitudes based on the niche models of different lineages. The map depicts lineage boundaries and the locality data corresponding to each latitudinal division. The panels to the left of the map show results for the parapatric boundary between the Coastal-Cascade and North-Central (NC) and *A. m. sigillatum* (sig) and Central-Oregon Highlands (COH) lineages. The panels to the right of the map show results for the parapatric boundary between the North-Central and Rocky Mountains (RM) and Central-Oregon Highlands and Rocky Mountains lineages. Contact zones in the comparisons are denoted with a star. In both sets of panels, the graphs to the left are based on the niche model of the western-most lineage in the comparison and the graphs to the right are based on the niche model of the eastern-most lineage in the comparison (also indicated by the colour of the average values in each plot). Plots are grouped to reflect the different ANOVA analyses (Table 3.3). Arrows at the top of the plots point away from the range of the focal lineage used to build the niche model.

Chapter 4: Patterns of introgression and hybrid performance in a northern contact zone between Long-Toed Salamander (Ambystoma Macrodactylum) lineages

4.1 Summary

Contact zones represent an opportunity to explore the processes influencing the maintenance of both genetic and geographic boundaries between divergent lineages. Although many studies have evaluated genetic patterns within contact zones, tests of the fitness of individuals in relation to these patterns are comparatively rare. In this study I take a multilocus approach to better characterize the genetic structure of a northern mitochondrial divide between two lineages of the long-toed salamander (*Ambystoma macrodactylum*). To evaluate hybrid performance in relation to this genetic divide, I assayed adult feeding performance (measured as prey consumed, total mass gained and food conversion efficiency) in a common garden experiment. The genetic results reveal a pattern consistent with the introgression of two markers, including mtDNA, from one lineage into what are otherwise genetically pure populations of the other lineage. Thus the transition between mitotypes appears to be removed from the average genomic transition between lineages. Intriguingly, feeding performance, especially for males, was lower in the region coinciding with the maximum extent of introgression and cytonuclear discordance than in the region demonstrating the most admixture. These findings—though suggestive of a possible role of feeding performance in limiting further introgression—were unexpected and underscore the potential complexity of patterns of genetic structure and variation in individual performance in hybrid zones.

4.2 Introduction

Species inhabiting previously glaciated regions are often characterized by marked genetic structure, forming a "patchwork of distinct genomes" (Hewitt 2004) on the landscape. This structure reflects, in part, the outcome of the sorting of genetic variation and evolution of genetic differences that took place during the Pleistocene as species underwent repeated episodes of range contraction and expansion (Hewitt 2000). At present, many of

these lineages encounter each other in contact zones where they may hybridize (Hewitt 2004). These contact zones not only serve as the range limits between groups but often contribute to broader suture zones that demarcate major biological provinces in the northern hemisphere (Swenson and Howard 2005; Rissler and Smith 2010). Thus contact zones represent an opportunity to study the processes influencing the integrity of genomes and the distribution of biodiversity alike.

Several outcomes are possible when closely related lineages hybridize. In the absence of any barriers to reproduction, lineages may exchange genes freely, leading to the formation of hybrid swarms and the breakdown of lineage boundaries (*e.g.* Wiens *et al*. 2006; Latch *et al.* 2011). Alternatively, selection against or in favour of hybrid genotypes can facilitate either the maintenance of genetic differences or the introgression of genes respectively. Whether range limits between distinct genetic groups are maintained in such cases depends on the fraction of the genome influenced by selection. Hybrid inviability, which affects the whole genome, can lead to sharp parapatric boundaries between lineages (Goldberg and Lande 2006). Less extreme selection against hybrids involving a number of loci can also maintain parapatric range limits between cohesive genetic groups (Barton 1983; Barton and Bengtsson 1986), manifest as concordant clines in allele frequency (*e.g.* tension zones: Barton and Hewitt 1985). Distinct boundaries between lineages can also be maintained when hybrids have higher fitness than parentals if this fitness advantage applies to a spatially restricted environment (*e.g.* bounded hybrid superiority: Moore 1977). The genetic structure of hybrid zones is expected to be more complicated when the direction and strength of selection varies across the genome or changes according to a patchy environment (Arnold 1997). Under these circumstances, hybrid zones may come to represent mosaics of different hybrid (Arnold 1997) and/or parental genotypes (Howard 1986; Harrison 1986, 1990). Thus, depending on the nature of selection, hybridization can have very different consequences for the genetic and geographic boundaries between lineages.

Many studies have characterized the genetic structure of contact zones and used these results along with cline theory to infer the nature of selection (more recent examples include Brelsford and Irwin 2009; Nolte *et al*. 2009; Singhal and Moritz 2012; Hamilton *et al*. 2013). Fewer studies have directly assessed the relative fitness of hybrid genotypes (see references in Arnold 1997 for examples as well as Parris 2000, Campbell and Waser 2007). Thus the

extent to which variation in fitness matches expectations based on observed genetic patterns in hybrid zones remains largely untested. Furthermore, much emphasis to date has been placed on the role of hybrid inviability or sterility in shaping hybrid zone dynamics. In cases where breeding adult populations are known, it may be important to consider the impact of other forms of hybrid performance (*e.g.* Sage *et al*. 1986; Fitzpatrick 2008), the cumulative effects of which may have important consequences for gene flow between lineages.

In the present study I explore the genetic structure of a northern contact zone between two lineages of the long-toed salamander (*Ambystoma macrodactylum*) and evaluate the feeding performance of individuals in relation to the genetic patterns observed. The long-toed salamander is a common pool-breeding amphibian in western North America (Figure 4.1). The species is currently divided into five subspecies on the basis of slight differences in average morphology (Ferguson 1961). Recent genetic studies have confirmed that the species consists of several distinct lineages (Figure 4.1; Thompson and Russell 2005; Savage 2008; Lee-Yaw and Irwin 2012), with estimates of mitochondrial divergence suggesting that much of this diversity arose during the Pleistocene (Savage 2008). My previous genetic survey across the northern portion of the species' range identified several places where these lineages are now in contact (Lee-Yaw and Irwin 2012). These areas represent an opportunity to examine the degree of reproductive isolation between the different long-toed salamander lineages and thus to address the consequences of secondary contact for genetic groups shaped by the Pleistocene glaciations.

Here I focus on the northernmost point of contact between the North-Central (NC) and Rocky-Mountains (RM) lineages (as described by Lee-Yaw and Irwin 2012; roughly concordant with *A.m columbianum* and *A. m. krausei* respectively based on Ferguson [1961]). Available estimates suggest that these lineages diverged during the late Pliocene (~3.7 mya; Savage 2008); but their current distributions indicate that they were restricted to separate glacial refugia during the Pleistocene. These lineages currently come into contact (based on the presence of both mtDNA lineages in the same pond or in close proximity) in several places along an extensive boundary associated with the Rocky Mountains (Figure 4.1; Lee-Yaw and Irwin 2012). I previously observed a transition between mtDNA hapolotypes where the two lineages meet along the Kicking Horse River in southeastern British Columbia (Lee-Yaw and Irwin 2012). The sharpness of this and other mtDNA clines

suggests that contact zones between the lineages are narrow, potentially reflecting strong barriers to gene flow. Nevertheless, the presence of several individuals carrying RM mtDNA but grouping more closely with NC individuals at nuclear markers (Lee-Yaw and Irwin 2012) suggests that hybridization has occurred in this area. However, the number of samples in that study was limited, restricting the conclusions that could be made about the molecular details of this northern contact zone. The first objective of the present study was thus to more fully evaluate genetic patterns at the nuclear genome using a large sample of individuals from this area.

My second objective was to evaluate the performance of individuals in relation to their genetic structure. I chose to focus on feeding performance (prey consumption and the ability to convert food energy to mass). Feeding performance affects the wellbeing of any organism and populations can show marked genetic differences in this trait (Billerbeck *et al.* 2000; Jonassen *et al.* 2000; Purchase and Brown 2000). For temperate amphibians such as the long-toed salamander, the ability to acquire resources during the summer may be particularly important for overall fitness. Many amphibians rely on fat stores built up during the summer feeding period for overwinter survival (*e.g.* Brenner 1969) and the production of gametes (Fitzpatrick 1976)—gamete maturation in long-toed salamanders specifically occurs during the winter months (Verrell 2006), when the animals are dormant and food resources are presumably limited. The ability to acquire food resources during the summer also influences body size in salamanders (Scott and Fore 1995), with larger individuals benefiting from greater fecundity (Kaplan and Salthe 1979; Scott and Fore 1995) as well as an advantage in mating (*e.g.* Chandler and Zamudio 2008) and other types of competitive interactions (*e.g.* Mathis 1990).

Because of these life history characteristics, it follows that hybrid dysfunction (or conversely, improved functionality) in feeding performance would have important consequences for fitness. Indeed, reductions in the ability of hybrids to obtain or process food is implicated in reproductive isolation in a number of systems where parental forms differ in traits associated with food resource acquisition (*e.g.* Darwin's Finches during some years: Grant and Grant 1996; crossbills: Benkman 1993; sticklebacks: Schluter 1995, but see Taylor *et al.* 2012). Although there is some evidence that NC and RM long-toed salamanders differ in vomerine tooth count (a means for securing prey items during ingestion; Ferguson 1961),

the extent to which they differ in other aspects of feeding morphology and ecology is unclear (see Chapter 3). However, reductions in hybrid feeding performance need not be associated with trait divergence. Both food consumption and conversion efficiency are mediated by a number of physiological processes and genetic incompatibilities at any of the loci responsible for these processes (*e.g*. Barendse *et al*. 2007) could presumably result in hybrid dysfunction (*i.e*. through Dobzhansky-Muller incompatibilities or duplication, degeneration and complementation: see Burke and Arnold [2001] for review). I thus chose to directly assay the feeding performance of individuals in relation to their genetic background.

Combined, the two aspects of my study address the following questions: 1) To what extent does the previously observed transition between mtDNA haplotypes reflect patterns in the nuclear genome? 2) Is there any variation among individuals in feeding performance? 3) If so, is this variation associated with the distribution of hybrid genotypes? My study thus attempts to link genetic variation to differences in performance within a northern hybrid zone. I report variation among markers in the extent of introgression and reduced feeding performance associated with the transition between some of these markers.

4.3 Methods

4.3.1 Animal collection and housing

Breeding adults (N=149) were captured during a two-week period in spring 2010 from eight ponds spanning the transition between NC and RM mitotypes in British Columbia and Alberta (Figure 4.1). Salamanders were housed individually in the laboratory in plastic shoe-box terrariums (34.9 x 20.3 x 12.7 cm^3) consisting of damp potting soil and a water dish that also served as a cover object. Animals were allowed to acclimatize to laboratory conditions for four weeks prior to the start of data collection. During this time, individuals were held at light-dark cycle of 14 h-19°C:10 h-14°C. Males spend the duration of the breeding season in ponds and feed during this period (Anderson 1968). In contrast, females do not appear to feed during the breeding season (Anderson 1968). To sustain male feeding and avoid extended starvation of the females during weeks when they would be leaving the ponds, individuals were fed five crickets per week during the acclimation period. Although arbitrary, this represents a modest amount of food (*e.g.* Scott and Fore 1995) and there is no reason to believe that individuals entered the experiment in different conditions as a result of

this treatment. All females finished depositing eggs during this time and well before the start of data collection.

4.3.2 Genetic data

I extracted DNA from tail clips of all individuals using a standard phenol-chloroform protocol. To assess the genetic ancestry of individuals, I scored individuals for three classes of genetic markers. First, to assess general levels of admixture, I generated AFLP profiles for all individuals in the experiment as well as 19 individuals (7 NC, 12 RM) from populations towards the centre of the range of each lineage. AFLPs were generated from four selective primer combinations (Table C1) following the protocols described by Lee-Yaw and Irwin (2012) and references therein. A replicate of each sample (protocol repeated from DNA extraction) was included in the dataset. Automatic binning and scoring of fragments was done in RawGeno v. 2.10.1 (Arrigo *et al*. 2012). After filtering out bins according to size (100-200 bp; chosen based on inspection of the electropherographs) and repeatability (>80%) criteria, the final AFLP dataset included 182 loci (average repeatability 91.4%). Differences between replicate PCRs were coded as missing data (approximately 7% of the final dataset, *i.e.* across all loci and all individuals). Six individuals from the study region were removed from the final AFLP dataset due to PCR error.

To assess the ancestry of individuals, I used the Bayesian clustering algorithm implemented in STRUCTURE (Pritchard *et al.* 2000; Falush *et al*. 2007) to calculate posterior probabilities of the assignment of individuals to 2 groups (an *a priori* expectation given the results of Lee-Yaw and Irwin [2012]; the analysis was also repeated for values of K up to 10 and K=2 was the optimal value according to the method of Evanno *et al*. 2005). STRUCTURE was run using the admixture model of ancestry. Ten runs of 1,000,000 MCMC generations were conducted with the first 500,000 generations discarded as burn-in. To explore variation among the AFLP loci in their contribution to any structure observed, I calculated marker-specific F_{ST} between the allopatric controls and between the two most widely separated populations within the transect. F_{ST} calculations were based on the method of Lynch and Milligan (1994) as implemented in AFLP-SURV v. 1.0 (Vekemans 2002; Vekemans *et al.* 2002). Allele frequencies were estimated in this analysis using a Bayesian

approach with a uniform prior distribution under the assumption of Hardy-Weinberg equilibrium.

To assess introgression at diagnostic markers, I focused on mtDNA and three nuclear protein-coding genes that also resolve monophyletic relationships among the long-toed salamander lineages (see Savage, 2008): Homeobox D11 (*hoxd11*), G2-I16 *A. mexicanum* cDNA (*g2i16*; Putta *et al*. 2004) and Collagen, type I, alpha-1 (*col1a1*; Voss *et al.* 2001; Weisrock *et al.* 2006). For mtDNA, I used the PCR-RFLP protocol described by Lee-Yaw and Irwin (2012) to score individuals as having either NC or RM mitotypes. I designed a similar protocol for the three nuclear markers. Specifically, I used the primers described by Savage (2008) to amplify and sequence these genes in 2-5 individuals from each of the ranges of NC and RM (outside of the study transect). Sequences for each gene were manually processed in BioEdit v. 7.0.5.3 (Hall 1999) and then aligned using Clustal (Thompson *et al.* 1994). I generated consensus sequences for each gene and used NEBcutter v. 2.0 (Vincze *et al.* 2003) to identify restriction enzymes that would differentially cut PCR products based on fixed SNPs between these NC and RM individuals. The ability of these enzymes to diagnose these lineages was verified on an additional 6-12 individuals from the ranges of each lineage prior to using them to assess the genotype (homozygous NC, heterozygous, or homozygous RM) of individuals in the present study. For *col1a1* it was possible to diagnose subspecies from PCR product size on an agarose gel due to a 522 bp deletion present in NC individuals (see also Savage, 2008; for full details of the protocols for all genes see Appendix C.1). Tail clips and all genetic data were collected after individual feeding performance (below) had been assessed.

4.3.2.1 Feeding data

I explored feeding performance and mass acquisition during an eight-week period from late June to late August (2010), corresponding to most of the natural summer foraging period for the region. Salamanders were separated into treatments of 14 h-25°C:10 h-20°C and 14 h-19°C:10 h-14°C approximating the average maximum summer temperatures at the western-most (Revelstoke, British Columbia) and eastern-most (Canmore, Alberta) ends of the study transect respectively. Individuals were randomly assigned to temperature treatment with the constraint that an equal number of males and females from each pond were placed in

both conditions. Temperature was gradually increased over the course of a week for animals in the warm treatment. Individual containers were arbitrarily shuffled within each environmental chamber three times a week throughout the study.

Individuals were fed seven gut-loaded crickets twice a week for the duration of the feeding study. Crickets were three weeks old with an average mass = 0.044 ± *sd* 0.017 g (based on measurements taken during the first two weeks of the experiment). The number of food items was based on the maximum reported intake of another ambystomid species fed *ad libitum in* the lab (Ducey and Heuer 1991). Prey availability may fluctuate or be more limited in the field (*e.g.* Jaeger 1979); thus this study speaks to differences in feeding performance under ideal conditions and is probably conservative in this regard. Salamanders were weighed prior to the first feeding of the study and then weekly for the eight week study. Individual differences in the number of food items consumed became apparent after the first two feedings and thus I also recorded the number of crickets consumed by each individual from the second week onwards. I report total change in mass over the course of the full study as well as mass conversion efficiency for each individual calculated as MCE=change in mass/(number of crickets consumed × average cricket mass). For MCE, I used the change in mass associated with the seven-week period for which food intake data was available and treated starting mass as a covariate.

4.3.2.2 Statistical analyses

Although some markers demonstrate transitions in allele frequency across the sampling transect, marker discordance combined with the low resolution of the AFLP dataset (see Results) made it difficult to classify individuals according to a hybrid index. However, at the pond-level, three broad groups emerge based on the amount of introgression of diagnostic markers (see Results; Figure 4.2). These genetic groups correspond with the natural grouping of ponds based on geographic proximity and thus for the remainder of the analyses I ask whether there are differences in the performance of individuals from western (Ponds 1-3; no introgression), middle (Ponds 4-6; introgression of RM mtDNA) and eastern (Ponds 7-8; introgression of both RM mtDNA and *g2i16*) sections of the transect.

The effects of genetic group, temperature and the interaction between genetic group and temperature on the different aspects of performance measured (number of food items

consumed, total weight gained and mass conversion efficiency) were examined using mixed-effects linear models in the *nlme* package in R (Pinheiro *et al*. 2011). In all models, pond of origin was treated as a random effect to deal with pseudoreplication arising from the spatial aggregation of individuals. The mass of individuals at the start of the experiment was included as a covariate in the models exploring differences in total weight gain and mass conversion efficiency. Because long-toed salamanders exhibit sexual dimorphism in at least some aspects of size (Ferguson 1961) and because selection to put on mass during the summer may differ between males and females, I ran the analyses for males and females separately. The response variables met the assumptions of normality and homogeneity of variance among treatments in all tests except for number of food items consumed, which was highly skewed and in the case of males, also violated the assumption of equal variances. Data transformations did not help in this case so I averaged food intake for each pond in each treatment and ran the models for food intake with pond as the unit of replication. This resulted in a dataset that better fit the assumptions of ANOVA.

4.4 Results

4.4.1 Extent of admixture and introgression across the zone

None of the AFLP loci surveyed demonstrated fixed differences between allopatric NC and RM individuals from outside of the study transect. However, 42 markers (23%) were found to have F_{ST} >0.01 (of which 7 have F_{ST} >0.1) between these allopatric controls. Results from the STRUCTURE analysis suggest that these data are sufficient for differentiating between NC and RM individuals as individuals from allopatric areas outside of the transect clearly fell out as two groups in that analysis (Figure 4.2). Surprisingly, the STRUCTURE analysis grouped all individuals across the study transect with NC individuals from other parts of the range with very little suggestion of RM ancestry (*e.g.* beyond background noise also present in the allopatric samples; Figure 4.2). Estimates of F_{ST} between the two most widely separated populations in the transect (Ponds 1 and 8) fell close to zero for most markers (Figure 4.3). Of those loci that do demonstrate F_{ST}>0.01 between these ponds, only two also differentiate the allopatric controls with F_{ST} >0.1 (Figure 4.3). Although the resolution of the AFLP dataset is low, these results suggest that individuals in the study transect derive most of their nuclear genome from NC. In contrast, although individuals from

the western section of the study transect possess NC mitotypes, all but one individual from the middle of and all individuals from the eastern section of the study transect possess RM mitotypes (Figure 4.2).

Results from the three diagnostic nuclear markers shed additional light on patterns of introgression and discordance in the region. Two of these markers, *col1a1* and *hoxd11*, largely corroborate the AFLP data. All individuals from the western section and middle of the transect were found to be homozygous for the NC allele at both genes, although the RM allele for both genes is present at the eastern edge of the transect/species' range (Figure 4.2). Interestingly, individuals that possess the RM allele are almost exclusively heterozygous at both of these loci (Figure 4.2). The pattern for *g2i16* differed markedly from both of these genes. In line with the pattern observed at the mtDNA, the western and eastern sections of the transect showed near fixation for the different *g2i16* alleles (Figure 4.2). Most individuals from the middle of the transect were homozygous for the NC *g2i16* allele, although the RM allele is present in these populations and is of equal frequency to the NC allele in Pond 4 (Figure 4.2).

I had too few diagnostic markers to score individuals according to a hybrid index using standard approaches (*e.g.* Buerkle 2005) and mixed ancestry was not apparent from the AFLP data. Nevertheless, examination of all markers combined supports the grouping of ponds into three broad genetic groups based on levels of introgression (Figure 4.2). Ponds 1-3 from the western section of the transect appear to represent NC populations, with all markers suggesting that individuals from these ponds derive majority of their genome from this lineage. RM mtDNA is found in ponds 4-6, thus defining a region where some introgression has occurred but where most individuals retain largely parental NC genotypes. Finally, ponds 7 and 8 from the eastern section of the transect show near fixation for RM alleles at two of the diagnostic markers and contain RM alleles at the remaining two diagnostic markers, thus defining the most RM-like region in the transect. This broad genetic grouping of ponds provides an intuitive framework for evaluating the results from the feeding experiment.

4.4.2 Feeding performance

Guided by the results above, I asked whether feeding performance differed among individuals from ponds that demonstrate different histories of introgression (hereafter referred to as genetic group). In terms of acquiring prey, the majority of individuals were able to consume all of the food items provided to them (Figure 4.4 a, b). Of those that did not, most were male. Genetic group was significantly associated with food intake among males, with males from the middle of the transect ingesting fewer food items than males from either end of the transect (linear mixed-effects model: $F_{df:2,5=}35.39$, p=0.001). In contrast, none of the females ate less than 89% of the food items provided (Figure 4.4 a) and there were no significant differences in food intake among genetic groups (Table 4.1). Neither temperature nor the interaction between temperature and genetic group had a significant effect on food intake for either sex (Table 4.1).

Considerable variation was observed in the total mass gained over the study for both females and males (Figure 4.4 c, d). Some of this variation is explained by differences in initial body size, as there was a significant negative relationship between initial mass and total weight gain for both sexes (Table 4.2). However, genetic group explained a significant proportion of the variation above and beyond the effect of body size for both sexes, with individuals from the middle of the transect gaining less mass than individuals from either end of the transect (Table 4.2, Figure 4.4). Once again, neither temperature nor the interaction between temperature and genetic group had a significant effect on weight gain (Table 4.2).

Patterns in mass conversion efficiency largely mirrored those observed for total mass gain (Figure 4.4 e, f). However, although the differences between genetic groups were significant in males, the differences between groups were not significant in females (Table 4.2). However, examination of the data revealed a single female in Pond 1 that had markedly lower conversion efficiency than all other individuals in this pond and one of the lowest values amongst females. This individual demonstrated an unusual growth trajectory, gaining the most weight of any salamander during the first week and then loosing weight again before a final gain. The association between genetic group and conversion efficiency was only marginally significant ($F_{df:2,dendf:5}$=5.74, p=0.051) after removing this outlier from the analysis. Removal of this individual from the analysis of total mass gained (above) did not qualitatively change those results.

4.5 Discussion

The present study was motivated by earlier results based on mtDNA and AFLP data from a limited number of individuals that suggested that the focal region is a zone of secondary contact between two long-toed salamander lineages (Lee-Yaw and Irwin 2012). I set out to further explore patterns of admixture and introgression in the region, expecting to find a general transition between NC and RM long-toed salamanders. Instead I observed a pattern more consistent with the introgression of a small number of RM markers (including mtDNA) into what are otherwise NC populations. Thus my study transect appears to capture the edge rather than the centre of the genomic transition between these two lineages. In this context, the reduced performance of individuals from the middle of the transect and not from the eastern edge of the transect (*i.e.* closer to the genome-wide centre of the hybrid zone) was unexpected and raises new questions about the factors influencing the introgression of different markers in this system.

4.5.1 Genetic characterization of the study region

The genetic data confirm a transition between mtDNA haplotypes across the study region. However, this transition is not representative of the average genomic transition between lineages. Instead, the clear grouping of individuals with NC individuals from other parts of the range based on the AFLP data suggests that the transect is composed of individuals deriving most of their nuclear genetic ancestry from NC. The fixation of two diagnostic nuclear markers for NC alleles in all but the eastern-most populations in the transect further supports this interpretation. However, some RM alleles are present in the region. I observed the near-fixation of RM *g2i16* alleles in the eastern section of the transect. Combined with the patterns observed for the mitochondrial genome and the observation of other RM alleles in the eastern section of the study transect, these data suggest that the transect is best described as the western periphery of the average genetic transition between NC and RM long-toed salamander lineages. Because my sampling coincided with the eastern range limit of *A. macrodactylum* and populations to the north fall within the range of NC (Lee-Yaw and Irwin 2012) the average genomic transition between NC and RM individuals is expected to fall to the south of the transect.

My results add to a growing number of studies documenting cytonuclear discordance (Toews and Brelsford 2012). Two alternative biogeographical scenarios can account for a mtDNA cline that is displaced from the nuclear transition between hybridizing species. The first scenario involves the movement of the hybrid zone such that a wake of mtDNA is left behind as the majority of the nuclear genome of one lineage advances into the range of the other—in the present case, this scenario would hold that NC has moved into the range of RM. At least two hypotheses have been put forth to explain this scenario. Wirtz (1999) discusses the potential for interspecific differences in the ability of males to entice matings from heterospecific females to lead to unidirectional hybridization and the presence of only one mtDNA lineage in hybrids. Krosby and Rohwer (2009) suggest that in the case of interspecific differences in male competitive ability, such unidirectional hybridization can lead to the mtDNA of the less aggressive species being left behind as the more competitive species expands its range. They argue that such aggressive hybridization explains the presence of Hermit warbler mtDNA well into the range of otherwise phenotypically pure Townsend's warblers in western North America (Krosby and Rohwer 2009).

Petit and Excoffier (2009)—based on the simulation results of Currat *et al.* (2008)—offer an alternative explanation for mitochondrial wakes following the expansion of one lineage into the range of the other. Their verbal model holds that because invading populations are small, alleles can surf to high frequencies via drift at the expansion front. If gene flow is high among populations of the invading lineage, alleles from this lineage are likely to be continuously introduced into the hybrid zone where they may surf to high frequencies (*i.e*. swamping alleles from the original resident lineage; Petit and Excoffier 2009). Sex-biased dispersal may serve to limit the introduction of sex-linked alleles from the invading species into the hybrid zone (Currat *et al.* 2008). Thus for sex-linked markers, such as mtDNA, alleles from the original resident species may persist at high frequencies as the invasion front moves through. Although it is possible to invoke either of these models and interpret the present results as reflecting the expansion of the range of NC at the expense of RM, the main assumption that NC individuals are competitively superior to RM individuals or have otherwise been able to take over this portion of their range requires further testing. Furthermore, it is unclear whether long-toed salamanders demonstrate sex-biased dispersal and thus whether the model of Petit and Excoffier (2009) could apply.

The alternative biogeographic scenario posits that RM mtDNA has moved into the range of NC. Several studies have reported selection on mtDNA variants (reviewed by Dowling *et al.* 2008) and theoretical work has revealed that even weak selection can lead to narrow phylogeographic breaks in mtDNA that do not reflect variation in nuclear markers (Irwin 2012). Thus one possible explanation for the patterns observed presently is that RM mtDNA affords a selective advantage to individuals and has advanced into the range of NC ahead of other alleles. Selection may similarly explain the somewhat more limited introgression of RM *g2i16* alleles in the eastern section of the transect. In general, under the hypothesis of adaptive introgression, only those markers under selection or closely linked to markers under selection are expected to demonstrate clines widely displaced from the centre of the hybrid zone (*e.g.* following from Maynard Smith and Haigh 1974; Barton 2000). This hypothesis thus offers a more parsimonious explanation for a few markers from one lineage being observed in individuals that otherwise have the genetic background of another lineage. Increasing the resolution of the nuclear dataset with more diagnostic markers and directly evaluating the effects of mitochondrial genotype on fitness (*e.g.* Flight *et al.* 2011) would help clarify the role of selection on the discordant patterns observed presently.

4.5.2 Reconciling feeding performance with observed genetic patterns

At the outset of this study, I expected to find a general transition between NC and RM genotypes in this region and a reduction in feeding performance at the centre of this transition. The genetic data suggest that the study transect is actually displaced from the average genomic transition between lineages. Yet intriguingly, animals from the middle of the transect gained less mass over the course of the summer than animals from either edge of the transect. Males from the middle of the transect also consumed fewer prey items and had poorer mass conversion efficiency. These results are not easily aligned with any of the explanations for the genetic patterns listed above nor by classic models of hybrid zones. For instance, performance is not expected to vary across the transect if the observed genetic patterns reflect the competitive take over of and subsequent introgression of NC alleles into part of RM's range. Likewise, under a hypothesis of adaptive introgression of mtDNA, individuals with RM mitotypes might be expected to demonstrate greater feeding performance than individuals with NC mitotypes, or at least perform equally as well if

feeding performance is decoupled from mtDNA function. Both the clear transitions in allele frequency and the patterns of performance observed would argue against describing this contact zone as either a case of bounded hybrid superiority or a mosaic hybrid zone. Yet if my sampling is simply off-centre with respect to what is largely a tension zone, I would expect a general decline in performance towards the eastern edge of the transect (*e.g.* towards the centre of the transition between forms) rather than a drop in performance in the middle of the transect. Thus results from the feeding experiment are perplexing in light of the genetic patterns observed.

One possible explanation for these results is that feeding performance is influenced by epistatic interactions and that hybrid individuals from the centre of the transect are lacking favourable combinations of alleles. Mitochondrial function serves as an illustrative example. In particular, the proper functioning of mtDNA depends on interactions with the nuclear genome (Dowling *et al.* 2008). If RM mtDNA works best with RM alleles for the genes in question and any of these genes demonstrate patterns similar to that observed for *g2i16*, mtDNA functionality would be expected to be low in populations from the middle of the transect and increase again at the eastern section of the transect. This explanation contradicts a hypothesis of adaptive introgression of RM mtDNA, although it is possible that RM mtDNA bestows individuals with a fitness advantage that outweighs the cost of any reductions in feeding performance. Regardless of whether mtDNA actually plays a role in shaping the patterns of feeding performance observed (see Bottje and Carstens 2009; Eya *et al.* 2011 for examples), the genetic results observed highlight the potential for there to be differences in the extent of introgression of different markers—some of which may be involved in epistatic interactions that influence feeding performance.

Alternatively, it is possible that feeding performance is shaped by environmental differences between populations. Although the feeding trials took place in a common garden, all individuals in the study were caught as adults from wild populations. Enivronmental conditions may vary across the transect and conditions experienced during earlier life stages influence adult feeding performance. Similarly, the patterns of feeding performance observed could be a bi-product of local adaptation to environmental differences (*e.g.* Purchase and Brown 2000) that happen to coincide with the area of admixture between lineages. Importantly, any environmental factor leading to reductions in feeding performance would

have to operate on multiple ponds from the centre of the transect (and not on ponds from either edge of the transect) and would therefore occur over fairly coarse spatial scales. Finally, it is noted that an environmental explanation for the observed performance of individuals may in turn tie into explanations for the extent of introgression in the system. Specifically, reductions in the performance of individuals owing to environmental conditions in the middle of the transect may cause populations in this area to function as demographic sinks. If the genetic patterns observed across the region reflect the movement of RM alleles into the range of NC, such sink populations could explain why introgression—adaptive or otherwise—has not proceeded past these populations.

4.6 Implications and future directions

It is typical for studies of hybrid zones to focus on reductions in hybrid performance at the centre of the average phenotypic or multilocus genetic cline between lineages. Less is known about the factors influencing the extent of introgression of markers that deviate from the average genomic pattern. Restricted by the limited availability of genetic data prior to commencing the study, I inadvertently (and fortuitously) assessed individual performance across a region corresponding to the periphery of the contact zone between two long-toed salamander lineages. The finding of reduced feeding performance coinciding with the extent of introgression and cytonuclear discordance serves as a potential example of an additional checkpoint—spatially removed from the front of hybridization—that helps maintain the boundary between genetic groups.

Nevertheless, several limitations of the present study need to be addressed before the observed patterns can be fully explained. Following from the earlier discussion, the role of the environment in shaping the observed patterns in feeding performance requires clarification. Repeating the experiment using individuals reared in the lab would shed light on the extent to which environmental conditions influence feeding performance. Likewise, conducting feeding assays under field conditions (see discussion in Arnold 1997) and extending the design to include a reciprocal transplant experiment may reveal important interactions between genotype and environment. In a related vein, data from additional populations spanning the genomic centre of secondary contact (presumably south of the transect) through to allopatric RM populations would help clarify the genetic and geographic

context of variation in feeding performance. This study was also limited in that I did not directly assess the effects of feeding performance on survival and fecundity (unfortunately attempts to subsequently breed the animals in the laboratory were unsuccessful). Although results from other ambystomid salamanders indicate that both the amount of prey consumed and body mass are related to fitness (*e.g.* Scott and Fore 1995), data specific to the long-toed salamander are necessary to address the ultimate demographic consequences of the patterns observed. Finally, exploring the mechanistic basis of differences in conversion efficiency, including the extent to which mtDNA is involved (*e.g*. Bottje and Carstens 2009; Eya *et al.* 2011) would allow for a more integrated explanation of the results (*e.g* Dalziel *et al*. 2009). This line of inquiry may include searching for other genes that correlate with feeding performance—a task that is becoming increasingly possible with the extension of genomic resources to non-model organisms.

4.7 Acknowledgements

I thank Chris Jacobs for the many hours of work he put in during the feeding trials and for his help with the collection of animals in the field. For help with field logistics, I thank Parks Canada (Revelstoke and Yoho National Parks), especially Lisa Larson and Derek Peterson. For animal care protocols, help and advice I thank Talia Sechley, Paul Verrell and Andrew Storfer. Mike Whitlock generously allowed me to occupy space in his environmental chambers for this experiment. Dolph Schluter provided important advice on the analyses. I thank Patricia Schulte and Gwylim Blackburn for many insightful comments on the results.

Table 4.1. Effects of temperature and genetic group (NC-like, admixed, RM-like) on food intake in a long-toed salamander contact zone[§].

			Female food intake		Male food intake	
Source	df	den df	F	P	F	p
Temperature	1	5	0.009	0.93	4.85	0.079
Genetic Group	2	5	1.27	0.36	35.39	0.0011*
Temperature* Genetic Group	2	5	0.37	0.71	4.529	0.075

[§]Sample location (pond) included as a random effect.

*p-value remains <0.05 after adjusting for multiple testing of the male dataset using the Benjamini and Hochberg (1995) step-up false discovery rate procedure available in the R multtest package (Pollard *et al*. 2005).

Table 4.2. Effects of temperature, genetic group (NC-like, admixed, RM-like) and starting mass on total mass gain and mass conversion effieciency (MCE) in a long-toed salamander contact zone[§].

	Females						Males					
			Mass Gained		MCE				Mass Gained		MCE	
Source	df	den df	F	p	F	p	df	den df	F	p	F	p
Temperature	1	38	1.76	0.19	0.97	0.33	1	87	2.25	0.14	0.57	0.45
Genetic Group	2	5	6.28	0.043	3.09 8	0.13	2	5	9.16	0.021*	9.05 0	0.022*
Starting Mass	1	38	15.04	0.0004*	1.22	0.28	1	87	12.44	0.0007*	8.49	0.0045*
Temperature x Genetic Group	2	38	0.33	0.72	0.58	0.56	2	87	0.55	0.58	1.29	0.28

[§]Sample location (pond) included as a random effect.

*p-values remain <0.05 after adjusting for multiple testing of the female and male datasets using the Benjamini and Hochberg (1995) step-up false discovery rate procedure available in the R multtest package (Pollard *et al*. 2005).

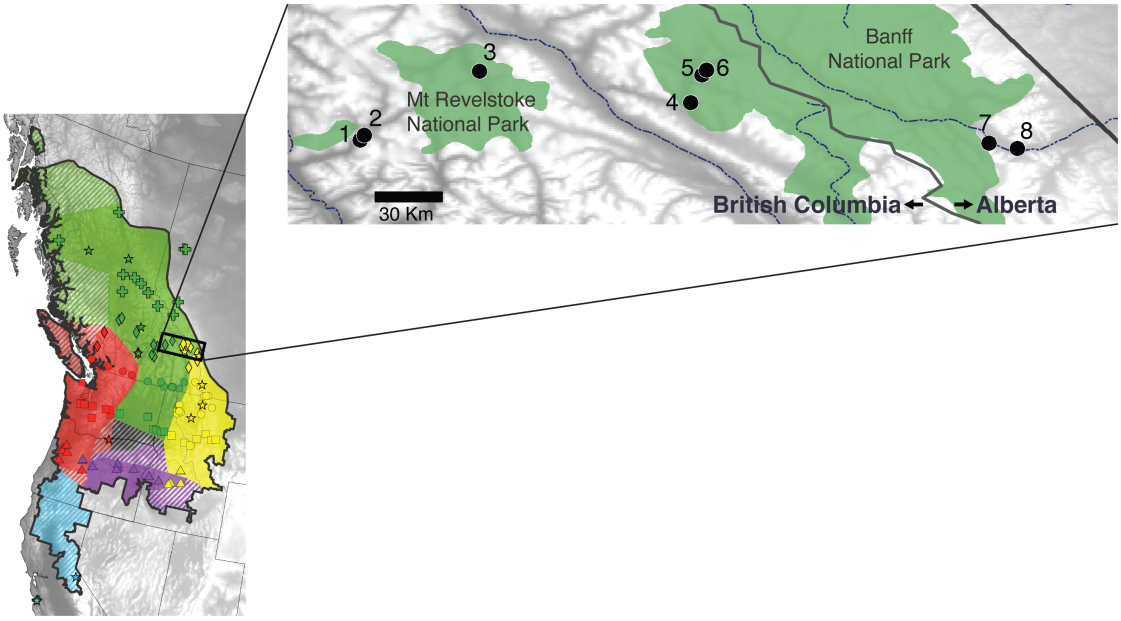

Figure 4.1. Breeding ponds from which adult long-toed salamanders were collected and position of sampling with respect to the distribution of the major genetic lineages identified by Lee-Yaw and Irwin (2012).

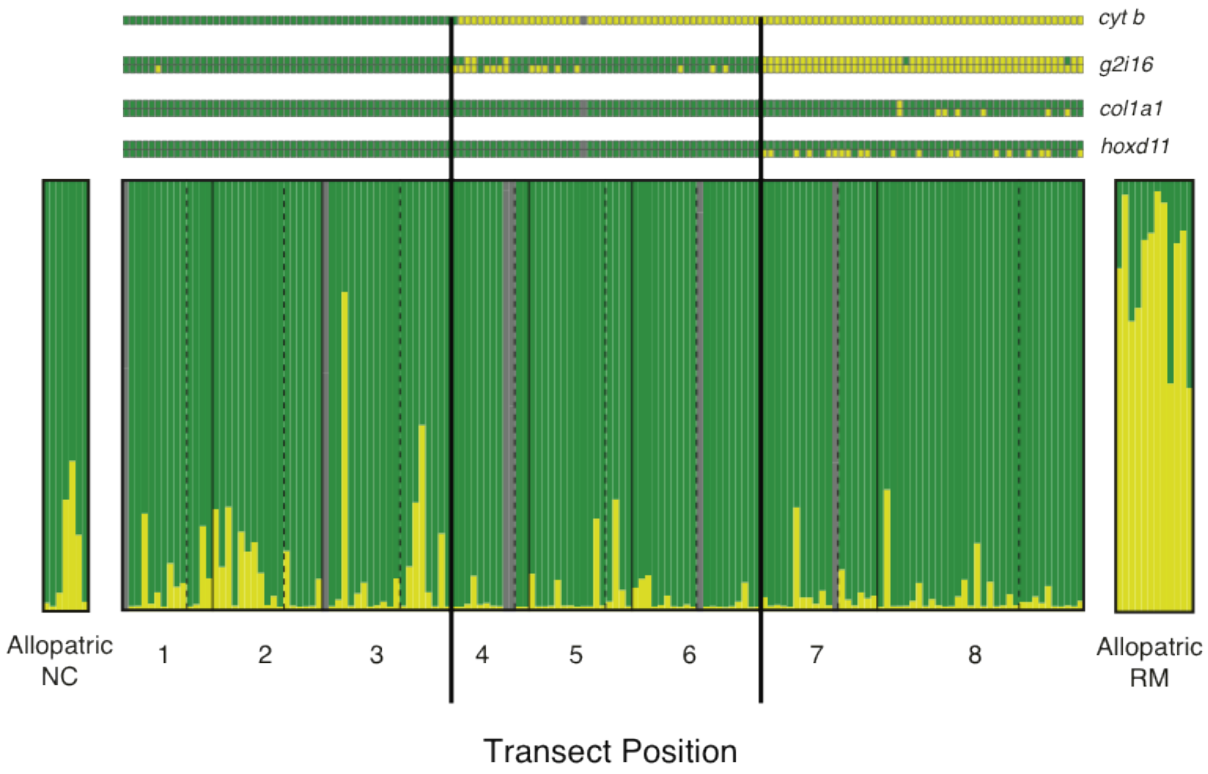

Figure 4.2. Genetic ancestry across a northern contact zone between North-Central and Rocky Mountains long-toed salamander lineages. The large bar plot shows the results from a STRUCTURE analysis of 182 AFLP loci, with bar height proportional to the posterior probability of assignment to the North Central (green) and Rocky Mountains lineages (yellow). Smaller bar plots above show the genotype of individuals at four diagnostic markers (mtDNA and three nuclear genes), with boxes representing the alleles carried by the individual at each marker. Grey boxes in all plots indicate missing data arising from PCR error. Individuals are grouped by pond (thin black lines) with pond numbers corresponding to those in Figure 4.1. Males and females within ponds are separated by a thin dashed line, with males presented first. The transect was divided into an eastern (Ponds 1-3), middle (Ponds 4-6) and western (Ponds 7-8) section (thick black lines) to assess differences in feeding performance.

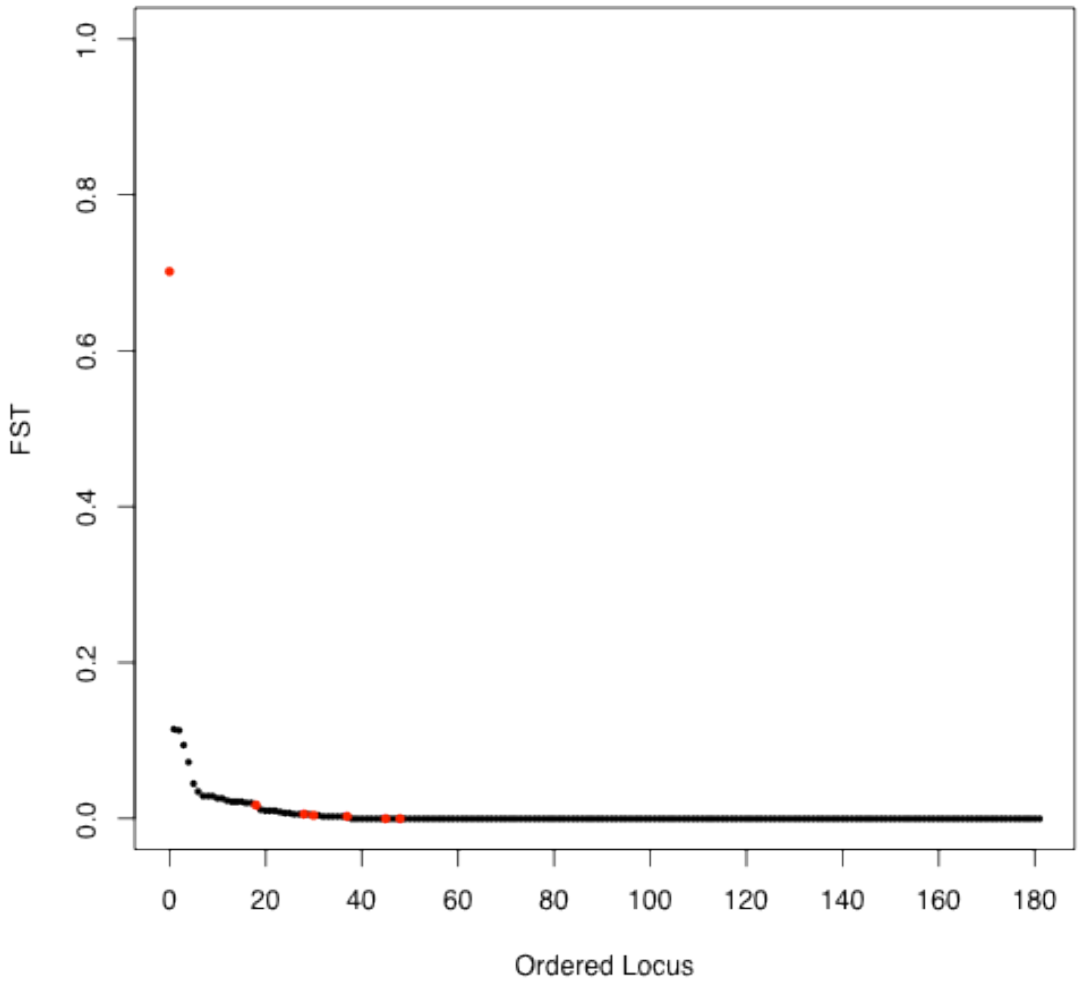

Figure 4.3. Locus-specific F_{ST} between edges of study transect (Ponds 1 and 8, see Fig. 4.1) based on AFLPs. Loci that demonstrate F_{ST}>0.1 between allopatric NC and allopatric RM individuals are coloured red.

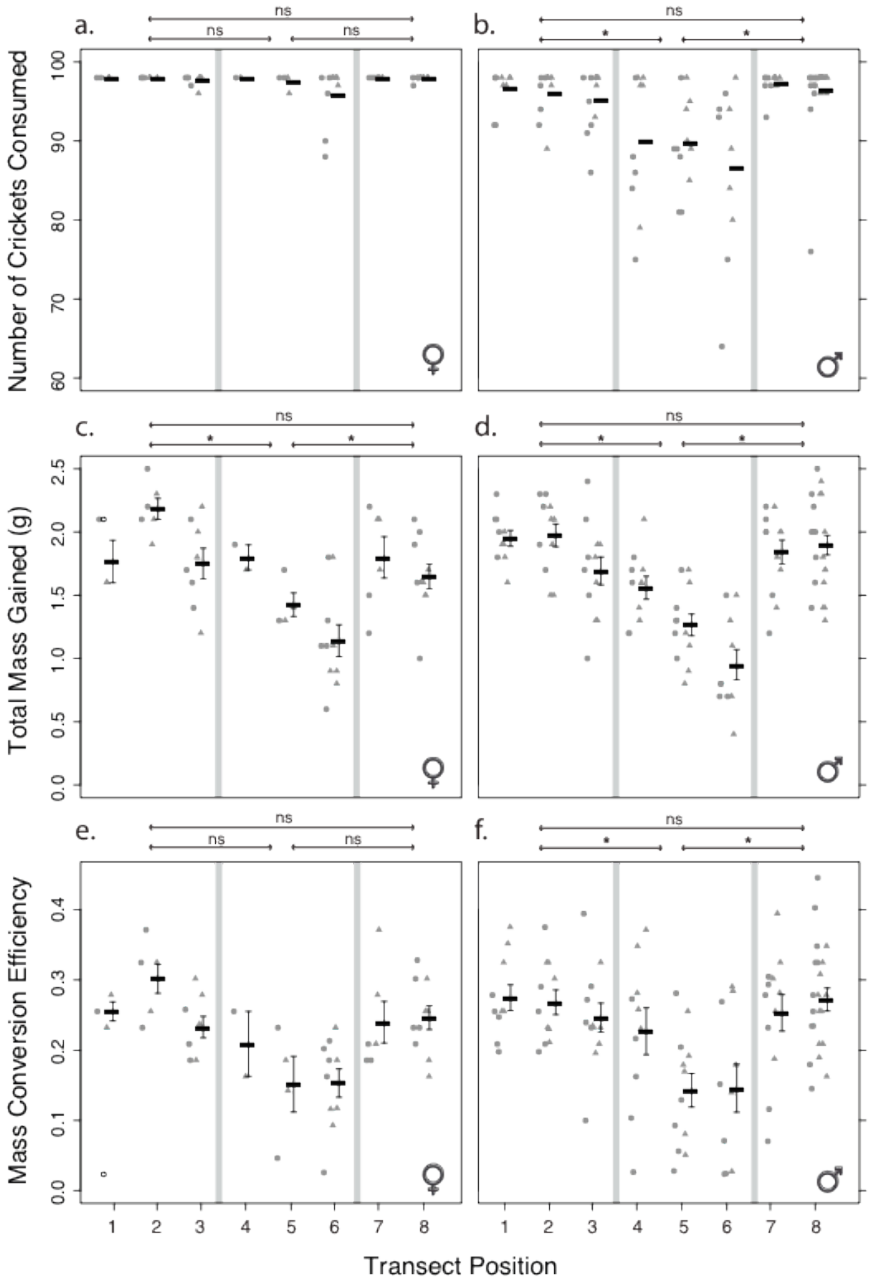

Figure 4.4. Feeding performance across a northern contact zone between NC and RM long-toed salamanders as measured by food intake (top panels), total mass gained during the summer (middle panels) and mass conversion efficiency (bottom panels). Results for females are shown on the left; results for males are shown on the right. Overall pond means are shown as thick black bars (with standard errors in the four bottom panels). Grey symbols denote the values for individuals and are differentiated according to temperature treatment (circles =19 ºC, triangles = 25 ºC). The one open circle in Pond 1 in panels c and e represents a single female who was an outlier with respect to growth trajectory (pond means for these plots do not include this individual). Significance at the α=0.05 level between groups of ponds with similar levels of introgression (see Figure 4.2) are reported above each plot, with thick lines separating ponds from these groups.

Chapter 5: General discussion and conclusions

Explaining species' geographic range limits is an outstanding challenge in evolutionary ecology and one that bears upon our ability to respond to some of the most imminent threats to biodiversity. Despite much attention (reviewed by Gaston 2003, Sexton *et al*. 2009; Gaston 2009), studies of the processes influencing species' range limits remain scattered across disparate fields and systems. The careful development of empirical systems and research programs that allow for tests of alternative hypotheses is sorely needed to better understand the relative importance of different range-limiting processes and to integrate diverse perspectives on the problem. Towards this goal, my dissertation focused on characterizing parapatric range limits and evaluating the relative importance of abiotic and biotic factors in shaping these range limits in a promising amphibian system—the long-toed salamander (*Ambystoma macrodactylum*). This work not only lays the foundation for further investigation into range limits in the long-toed salamander, but also illustrates general steps that can be taken when studying parapatric range limits in a novel system.

5.1 Major findings and implications of the research

5.1.1 The large geographic distribution of the Long-Toed Salamander consists of multiple divergent genetic lineages (Chapter 2)

Accurately documenting parapatric range limits is an important first step towards understanding them. Previous authors have proposed that the long-toed salamander is comprised of several distinct subspecies (Ferguson 1961; Thompson and Russell 2005; Savage 2008). However, descriptions of these groups have varied considerably, reflecting differences between various morphological and genetic markers. My genome-wide survey of genetic variation provides support for the existence of at least four distinct lineages of long-toed salamander, clarifying both the amount and distribution of diversity in this widespread species. Notably, my results highlight several places where existing subspecies boundaries fail to capture major genetic breaks in the system, including evidence for a cryptic lineage in the highlands of central Oregon.

These results contribute to a growing number of studies documenting the presence of highly divergent, cryptic lineages within what are currently treated as single taxonomic entities (reviewed by Bickford *et al.* 2006). In the case of the long-toed salamander, genome-wide differences between lineages argue for their separate treatment in subsequent work, thus confirming that the system represents an opportunity to undertake comparative studies of the factors influencing parapatric range limits in closely related pairs of taxa. However, more generally these results underscore the importance of assessing phylogenetic patterns prior to addressing range limit hypotheses in any system. For instance, tests of proposed distributional patterns (*e.g*. abundant centre: reviewed by Sagarin and Gaines 2002) and many of the hypotheses for range limits (*e.g*. limited genetic variation in peripheral populations: reviewed by Eckert *et al.* 2008; swamping gene-flow: Kirkpatrick and Barton 1997) involve making comparisons between central and peripheral populations. The presence of divergent genetic groups across a species' range—regardless of whether such groups are the focus of study—suggests that it may be important to consider the distinct evolutionary trajectories of these groups when designating central and peripheral populations and choosing sites for further study.

Even when phylogeographic information is available for a given system, the boundaries between major genetic groups are often described on the basis of just a handful of genetic markers. The potential for gene trees to deviate from the history of population divergence (Knowles and Maddison 2002) and for different markers to demonstrate substantial discordance is now widely recognized (Toews and Brelsfeld 2012). The increasing ease with which genomic data can be collected has made it possible to revisit existing single (or few) gene phylogenies to confirm genetic patterns. Nevertheless, genomic techniques have only recently become available for most non-model organisms and thus the extent to which phylogenies based on a limited number of markers represent genome-wide patterns remains largely unclear. My phylogeographic survey of the long-toed salamander, contributes one example towards addressing this question. In the case of the long-toed salamander, most markers diagnose the major genetic lineages. However, evidence for cryptic lineages as well as cytonuclear discordance along range edges, suggests that there is utility in revisiting existing phylogenies with genomic data when the goal is to accurately characterize the boundaries between genetic groups. In chapter 3, I demonstrate how such

data in combination with spatial analyses, can allow investigators to move beyond qualitative descriptions of genetic boundaries to more quantitative estimates of these boundaries.

5.1.2 Range limits of the Long-Toed Salamander lineages do not reflect the edge of their respective climatic niches (Chapter 3)

The extent of ecological divergence between parapatric taxa and the distribution of a shared border with respect to the fundamental niche of each species provides insight into the relative importance of different hypotheses for parapatric range limits. I used spatial data and ecological niche modeling to evaluate the influence of climate on patterns of diversity in the long-toed salamander. I specifically asked whether range limits can be explained by niche divergence and the boundaries of each lineage's respective climatic niche (*e.g.* Rissler and Apodaca 2007; Glor and Warren 2008) or by niche conservatism and the presence of climatic barriers separating otherwise suitable habitat for each lineage (Glor and Warren 2008).

I found that the different lineages of long-toed salamander are largely ecologically exchangeable in terms of the climatic space that they occupy. Furthermore, models of the climatic niches of the different lineages suggest that suitable climatic conditions for each lineage can be found beyond their current range limits. These results refute the importance of adaptation to different climatic conditions for contemporary range limits in this system and add to the growing number of studies showing poor correspondence between range limits and climate (Svenning and Skov 2004; Munguía *et al.* 2012). These results also speak to the ongoing debate as to the importance of ecology—and climate specifically (Keller and Seehausen 2012)—in driving lineage diversification, providing an example of divergence despite niche conservatism.

However, my results do not completely rule out a role for climate in shaping range limits in this system. Several of the boundaries between lineages coincide with areas where the average climatic suitability for either lineage is low. This pattern was observed across different sections of the same boundary in a number of cases, highlighting the potential for climate to consistently influence the boundary between a given pair of taxa. These results add to the growing number of studies suggesting that climatic conditions may serve to limit the extent of overlap (or lack thereof) between closely related taxa (Kozak and Wiens 2006; Rissler and Apodaca 2007; Glor and Warren 2008; Edwards *et al.* 2013; Soto-Centeno *et al*.

2013; Werner *et al*. 2013). However, I found that the suitability scores of sites within contact zones varied considerably even as average suitability declined in these areas. Such variability highlights the potential for climatic barriers to be "leaky", raising questions as to whether climatic barriers alone are sufficient for maintaining parapatric boundaries between taxa that demonstrate niche conservatism.

I note that major contribution of my work in Chapter 3 is the approach taken. Specifically, my analysis of the climatic niches of the different lineages incorporates many recent methodological developments in the field of niche modeling (*e.g*. randomization test of AUC values: Raes and ter Steege 2007; limitation of background extent: Anderson and Raza 2010; use of bias grids: Elith *et al*. 2010; use of MESS surfaces: Elith *et al*. 2010). In addition, my work brings together existing frameworks for evaluating the role of climate in shaping parapatric boundaries (*e.g*. Rissler and Apodaca 2007; Glor and Warren 2010; McCormack *et al.* 2010) and extends them by assessing spatial variation in the relationship between range limits and climatic suitability. Thus my work provides a strong methodological example of using ecological niche models to evaluate the role of climate in shaping range limits.

5.1.3 The transition between genetic groups varies widely among markers at fine spatial scales (Chapter 4)

The results above suggest that biotic interactions may be a more important driver of range limits in the long-toed salamander than abiotic factors. In light of evidence of hybridization between lineages (Chapter 2), my final goal was to evaluate the consequences of hybridization for the maintenance of lineage boundaries. Focusing on the northern-most contact zone between two of the long-toed salamander lineages, I conducted additional sampling and employed several markers to better characterize the genetic outcome of secondary contact. I observed clear transitions between these lineages for two of the diagnostic markers included in my study, including the mitochondrial marker. However, in contrast to either a tension zone model (Barton or Hewitt 1985) or models predicting sharp boundaries between lineages (Goldberg and Lande 2006), these clines are displaced with respect to patterns observed at other markers.

A growing number of studies have reported similar cytonuclear discordance in hybrid zones (reviewed by Toews and Brelsford 2012). More generally, the potential for genomes to be "porous" and for different markers to demonstrate different degrees of introgression depending on selection (reviewed by Wu 2001), is now widely acknowledged. These studies once again emphasize the importance of using multiple markers to characterize range limits when delineating morphologically cryptic taxa. More fundamentally, these studies highlight the potential for range limits to be "blurry" in cases where lineages hybridize. In the age of genomics, a gene-centric approach (*e.g.* Wu 2001) to the study of factors shaping parapatric range limits will likely prove necessary.

5.1.4 Feeding performance varies within a northern contact zone between two lineages of Long-Toed Salamander (Chapter 4)

Almost all range-limit hypotheses that involve hybridization hold that hybrid individuals are less fit than parentals (*e.g.* tension zones: Barton and Hewitt 1985; reproductive character displacement interacting with gene flow: Goldberg and Lande 2006; but see Moore 1977). My research provides one of few empirical tests of hybrid dysfunction in amphibians (see also Nürnberger *et al.* 1995; Parris 2001; Lemmon and Lemmon 2010). I specifically examined the feeding performance (measured as number of prey ingested, mass gained and conversion efficiency) of adult individuals from the above contact zone in relation to their genetic background. My initial intention was to assay the performance of individuals across the entirety of the contact zone; however, at the outset of my study very limited genetic information was available to guide my sampling and I inadvertently ended up working with individuals from what appears to be the western periphery of this contact zone.

In light of these genetic patterns, results from my feeding experiment were intriguing. I observed that feeding performance was lowest in the region coinciding with the maximum extent of introgression and cytonuclear discordance rather than in the region demonstrating the most admixture. Although my study could not distinguish between environmental or genetic explanations for this finding, these results may be relevant for understanding the limits to introgression. Specifically, the reduced feeding performance of individuals in this part of the contact zone may cause these populations to act as demographic sinks, limiting the subsequent spread of alleles moving across the zone.

Whether this hypothesis turns out to be correct, this work has important implications for studying fitness in hybrid zones. To date many genetic surveys of contact zones are based on just a few markers. Allelic transitions between groups for two of my markers coincide with the area of reduced feeding performance. Had I just had information from these markers, I may have erroneously concluded that I was dealing with a standard tension zone (*i.e.* reduced hybrid performance at the centre of the genomic transition between lineages: Barton and Hewitt 1985). My results therefore suggest that it may be important to revisit earlier studies of hybrid zones with genomic data in order to evaluate the true correspondence of studies with this leading hybrid zone model. In a related vein, Arnold (1997) has noted the potential for the fitness of hybrids to vary drastically depending on genotype and/or environment. My analysis of the periphery of a contact zone demonstrates the potential for this complexity to play out across different regions of a contact zone. The possible implications of these results for understanding the limits to introgression suggest that future studies of hybrid fitness may want to broaden their spatial scope to incorporate the full distribution of hybrid genotypes on the landscape.

5.2 Limitations of the research and future directions

5.2.1 Recommendations for further phylogeographic study

My work provides an extensive survey of genetic variation across the range of the long-toed salamander (Chapter 2). Although my sampling and choice of markers allowed for a genome-wide assessment of levels of divergence between the different lineages across most of the range of the species (Chapter 2), I note that additional work is needed to fill several gaps in my sampling. Most critically, I had very few samples from the ranges of the two putative long-toed salamander subspecies in California. Although the mitochondrial data that I do have, in combination with the work of Savage (2008), provides some support for the recognition of two distinct lineages in this region, additional samples and genomic data are necessary to verify these conclusions as well as delineate the boundaries of the lineage in the Sierra Mountains. More generally, my work in Chapter 2 and Chapter 3 is perhaps best viewed as a broad-scale estimate of genetic boundaries in this system. Now that we have a good approximation of the boundaries between lineages, additional fine-scale surveys are necessary to clarify the transition between lineages in areas that are currently under-surveyed

(*e.g.* Central Basin, Washington State) and to further elucidate the genetic details of contact zones in the system (*i.e.* Chapter 4).

Apart from increased sampling to further characterize the boundaries between lineages, our understanding of the biogeographic history of the species would benefit from additional phylogeographic analyses. For instance, genotyping more individuals from each population to assess patterns of genetic variation across the range could help verify the location of glacial refugia (*i.e*. identified as regions of high diversity, see Hewitt 1996). Likewise, sequence data from additional samples and markers could allow for the reconstruction of specific routes taken during post-glacial recolonization, thus speaking to the effects of different landscape features on range expansion (*e.g*. Hewitt 1999; Austin *et al.* 2001; Lee-Yaw *et al*. 2008). Combined with coalescence analyses, such data could also reveal the timing of different colonization events (*e.g*. Peters *et al.* 2008; McCracken *et al.* 2013), which may in turn help explain the dramatic differences in the northern extent of the different lineages (*e.g*. through priority effects: reviewed by Sexton *et al.* 2009).

5.2.2 Recommendations for further assessment of the abiotic niche

My work in Chapter 3 suggests that climate has not played a major role in the differentiation and distribution of the different long-toed salamander lineages. However, the methodology that I used represents an indirect approach to characterizing species' climatic niches. Others have argued for a more mechanistic approach to niche modeling, whereby information about the physiological tolerances of individuals is used to evaluate their potential fitness across the landscape (Kearney and Porter 2009; Buckley *et al*. 2010). These approaches are data- and labor- intensive, making it difficult to incorporate the collective effects of many variables into models of species' potential distributions. Nevertheless, evaluating the physiological tolerances of individuals from different lineages in response to key climatic variables could help verify my results. For instance, results from my niche models suggest that minimum winter temperature and summer heat relative to moisture are important predictors of the occurrence of most lineages on the landscape. Direct assessment of the response of individuals from different lineages with respect to these variables would thus seem pertinent to confirm the patterns of niche conservatism implied by my results.

In order to fully understand the extent of niche conservatism versus divergence in the system, further work is also required to determine whether the different lineages of long-toed salamander have diverged with respect to other aspects of the niche. For instance, amphibians are known to demonstrate local adaptation and divergence with respect to different vegetation, pond characteristics and food resources (see examples in Smith and Skulason 1996; Takahashi and Parris 2008). It is possible that range limits in the long-toed salamander reflect differences in these or other variables that were not included in my niche models. This possibility highlights the need to collect more basic information about the ecology of the different lineages before asserting that they are fully ecologically exchangeable. With the boundaries between lineages now clearly delineated, designing studies to collect this data should be relatively straightforward.

5.2.3 Recommendations for further study of the outcome of secondary contact

In my final data chapter, I explored both the fine-scale genetic structure of a contact zone between two long-toed salamander lineages and variation in feeding performance in relation to this structure. I found evidence of intrinsic differences in feeding performance in some areas of the contact zone. Although these results may have implications for understanding the extent of introgression in the region, I discussed several limitations to this study in Chapter 4. Here I note more generally that this work is just the beginning of understanding how hybridization and reproductive isolation shape the boundaries between lineages in this system. Future work characterizing the genetic structure of other contact zones will go a long way towards understanding the extent to which the genomic integrity of each lineage is maintained upon secondary contact. Likewise, tests of feeding performance in other contact zones are necessary to determine whether variation in this aspect of individual vigor is consistently associated with the transition between mitochondrial groups. That the most genetically admixed individuals performed as well as those with parental genotypes in my study suggests that other forms of hybrid dysfunction, if any, are more important general barriers to gene exchange in this system. Identifying the nature and number of these barriers is necessary to fully explain the maintenance of the genetic boundaries between lineages.

REFERENCES

Alleaume-Benharira M, Pen IR, Ronce O (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *Journal of Evolutionary Biology*, **19**, 203-215.

Ålund M, Immler S, Rice AM, Qvarnstrom A (2013) Low fertility of wild hybrid male flycatchers despite recent divergence. *Biology Letters*, **9**, 20130169.

Anderson JD (1967) A comparison of the life histories of coastal and montane populations of *Ambystoma macrodactylum* in California. *American Midland Naturalist*, **77**, 323-355.

Anderson JD (1968) A comparison of the food habits of *Ambystoma macrodactylum sigillatum*, *Ambystoma macrodactylum croceum* and *Ambystoma tigrinum californiense*. *Herpetologica*, **24**, 273-284.

Anderson JD (1972) Embryonic temperature tolerance and rate of development in some salamanders of the genus Ambystoma. *Herpetologica*, **28**, 126-130.

Anderson RFV (1977) Ethological isolation and competition of allospecies in secondary contact. *American Naturalist*, **111**, 939-949.

Anderson RP, Raza A (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, **37**, 1378-1393.

Angert AL, Schemske DW (2005) The evolution of species' distributions: reciprocal transplants across the elevation ranges of *Mimulus cardinalis* and *M. lewisii*. *Evolution*, **59**, 1671-1684.

Araújo MB, Pearson RG (2005) Equilibrium of species' distributions with climate. *Ecography*, **28**, 693-695.

Arif S, Adams, DC, Wicknick JA (2007) Bioclimatic modelling, morphology, and behaviour reveal alternative mechanisms regulating the distribution of two parapatric salamander species. *Evolutionary Ecology Research*, **9**, 843-854.

Arnason U, Gullberg A, Janke A, Joss J, Elmerot C (2004) Mitogenomic analysis of deep gnathostome divergences: a fish is a fish. *Gene*, **333**, 61-70.

Arnold ML (1997) *Natural hybridization and evolution*. Oxford University Press, New York.

Arnold ML, Hodges SA (1995) Are natural hybrids fit or unfit relative to their parents? *Trends in Ecology and Evolution*, **10**, 67-71.

Arrigo N, Holderegger R, Alvarez N (2012) Automated scoring of AFLPs using RAWGENO v 2.0, a free R CRAN library. In: *Data Production and Analysis in Population Genomics* (eds Pompanon F & Bonin A), pp. 155–176. Humana Press, Berlin.

Arteaga MC, McCormack JE, Eguiarte LE, Medellin RA (2011) Genetic admixture in multidimensional environmental space: asymmetrical niche similarity promotes gene flow in armadillos (*Dasypus novemcinctus*). *Evolution*, **65**, 2470-2480.

Austin JD, Lougheed SC, Neidrauer L, Chek AA, Boag PT (2001) Cryptic lineages in a small frog: the post-glacial history of the spring peeper, *Pseudacris crucifer* (Anura: Hylidae). *Molecular Phylogenetics and Evolution*, **25**, 316-329.

Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure (2013) Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews*, **88**, 310-326.

Barber BR, Jensen G (2012) Quaternary climate change was not an engine of diversification in New World bats (Chiropetera). *Journal of Mammalian Evolution*, **19**, 129-133.

Barendse W, Reverter A, Bunch RJ, Harrison BE, Barris W, Thomas MB (2007) A validated whole-genome association study of efficient food conversion in cattle. *Genetics*, **176**, 1893-1905.

Barton N (2001) Adaptation at the edge of a species' range. In: *Integrating Ecology and Evolution in a Spatial Context* (eds Silverton J, Antonovics J), pp. 365-392. Blackwell, London.

Barton NH (1983) Multilocus clines. *Evolution*, **37**, 454-471.

Barton NH (2000) Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London B,* **355**, 1553-1562.

Barton NH, Bengtsson BO (1986) The barrier to genetic exchange between hybridizing populations. *Heredity*, **56**, 357-376.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113-148.

Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Sóberon J, Villalobos F (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810-1819.

Baumann H, Conover DO (2011) Adaptation to climate change: contrasting patterns of thermal-reaction-norm evolution in Pacific versus Atlantic silversides. *Proceedings of the Royal Society of London B*, **278**, 2265-2273.

Becker RA, Chambers JM and Wilks AR (1988) *The New S Language*. Wadsworth and Brooks/Cole, Pacific Grove CA.

Benkman CW (1993) Adaptation to single resources and the evolution of crossbill (*Loxia*) diversity. *Ecological Monographs*, **63**, 305-325.

Berggren KT, Ellegren H, Hewitt GM, Seddon JM (2005) Understanding the phylogeographic patterns of European hedgehogs *Erinaceus concolor* and *E europaeus* using the MHC. *Heredity*, **95**, 84-90.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2006) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, **22**, 148-155.

Billerbeck JM, Schultz ET, Conover DO (2000) Adaptive variation in energy acquisition and allocation among latitudinal populations of the Atlantic Siverside. *Oecologica*, **122**, 210-219.

Blows MW, Hoffmann AA (2005) A reassessment of genetic limits to evolutionary change. *Ecology*, **86**, 1371-1384.

Bottje WG, Carstens GE (2009) Association of mitochondrial function and feed efficiency in poultry and livestock species. *Journal of Animal Science*, **87**, E48-E63.

Brenner FJ (1969) The role of temperature and fat deposition in hibernation and reproduction in two species of frogs. *Herpetologica*, **25**, 105-113.

Bridle JR, Vines TH (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology and Evolution*, **22**, 140-147.

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001) Comparative phylogeography of northwestern North America: a synthesis. In: *Integrating Ecological and Evolutionary Processes in a Spatial Context* (eds Silverton J, Antonovics J), pp 319-339. Blackwell Science, Oxford.

Buckley LB, Urban MC, Angilletta MJ, Crozier LG, Rissler LJ, Sears MW (2010) Can mechanism inform species' distribution models? *Ecology Letters*, **13**, 1041-1054.

Buerkle AC (2005) Maximum-likelihood estimation of a hybrid index based on molecular markers. *Molecular Ecology Notes*, **5**, 684-687.

Bull CM (1991) Ecology of parapatric distributions. *Annual Review of Ecology and Systematics*, **22**, 19-36.

Bullock JM, Edwards RJ, Carey PD, Rose RJ (2000) Geographical separation of two *Ulex* species at three spatial scales: does competition limit species' ranges? *Ecography*, **23**, 257-271.

Burke JM, Arnold ML (2007) Genetics and the fitness of hybrids. *Annual Review of Genetics*, **35**, 31-52.

Caccone A, Milinkovitch MC, Sbordoni V, Powell JR (1997) Mitochondrial DNA rates and biogeography in European newts (Genus *Euproctus*). *Systematic Biology*, **46**, 126-144.

Cambefort Y (1994) Body-size abundance and geographical-distribution of Afrotropical dung beetles (Coleoptera Scarabaeidae). *Acta Oecologica,* **15**, 165-179.

Campbell DR, Waser NM (2007) Evolutionary dynamics of an *Ipomopsis* hybrid zone: confronting models with lifetime fitness data. *American Naturalist*, **169**, 298-310.

Carstens BC, Degenhardt JD, Stevenson AL, Sullivan J (2005) Accounting for coalescent stochasticity in testing phylogeographical hypotheses: modeling Pleistocene population structure in the Idaho giant salamander *Dicamptodon aterrimus*. *Molecular Ecology*, **14**, 255-265.

Carstens BC, Stevenson AL, Degenhardt JD, Sullivan J (2004) Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. *Systematic Biology*, **53**, 781-792.

Case TJ, Taper ML (2000) Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. *American Naturalist*, **155**, 583-605.

Chandler CH, Zamudio KR (2008) Reproductive success by large, closely related males facilitated by sperm storage in an aggregate breeding amphibian. *Molecular Ecol*ogy, **17**, 1564-1576.

Chase JM, Leibold, MA (2003) *Ecological niches: linking classical and contemporary approaches*. The University of Chicago Press, London.

Chunco AJ, Jobe T, Pfennig KS (2012) Why do species co-Occur? A test of alternative hypotheses describing abiotic differences in sympatry versus allopatry using spadefoot toads. PLOS One. doi:10.1371/journal.pone.0032748

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1660.

Connell JH (1961) The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. *Ecology*, **42**, 710-723.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates Inc., Sunderland MA.

Cunningham HR (2009) Competition at the range boundary in the slimy salamander: using reciprocal transplants for studies on the role of biotic interactions in spatial distributions. *Journal of Animal Ecology*, **78**, 52-62.

Cunningham HR, Rissler LJ, Apodaca JJ (2009) Competition at the range boundary in the slimy salamander: using reciprocal transplants for studies on the role of biotic interactions in spatial distributions. *Journal of Animal Ecology*, **78**, 52-62.

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908-1920.

Dalziel AC, Rogers SM, Schulte PM (2009) Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. *Molecular Ecology*, **18**, 4997-5017.

Darwin CR (1859) *On The Origin of Species by Means of Natural Selection, or The Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Davies TJ, Meiri S, Barraclough TG, Gittleman JL (2007) Species co-existence and character divergence across carnivores. *Ecology Letters*, **10**, 146-152.

Davies TJ, Purvis A, Gittleman JL (2009) Quaternary climate change and the geographic ranges of mammals. *American Naturalist*, **174**, 297-307.

Dawes-Gromadzki TZ, Bull CM (1997) Laboratory studies of ant predation on parapatric reptile ticks. *Australian Journal of Ecology*, **22**, 1-7.

Dowling DK, Friberg U, Lindell J (2008) Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends in Ecology and Evolution*, **23**, 546-554.

Ducey, PK, Heuer J (1991) Effects of food availability on intraspecific aggression in salamanders of the genus Ambystoma. *Canadian Journal of Zoology*, **69**, 288-290.

Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4, 359-361. http://taylor0.biology.ucla.edu/structureHarvester/

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographic ranges: the central-marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170-1188.

Edwards RD, Crisp MD, Cook LG (2013) Niche differentiation and spatial partitioning in the evolution of two Australian monsoon tropical tree species. *Journal of Biogeography*, **40**, 559-569.

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330-342.

Elith J, Phillips SJ, Hastie T, Dudík M (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43-57.

Endler JA (1977) *Geographic variation, speciation and clines*. Princeton University Press, Princeton.

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

Excoffier L, Laval G, Schneider S (2005) Arlequin ver 30: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1**, 47-50.

Eya JC, Ashame MF, Pomeroy CF (2011) Association of mitochondrial function with feed efficiency in rainbow trout: Diets and family effects. *Aquaculture*, **321**, 71-84.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574-578.

Ferguson DE (1961) The geographic variation of *Ambystoma macrodactylum* Baird with the description of two new subspecies. *American Midland Naturalist*, **65**, 311-338.

Filin I, Holt RD, Barfield M (2008) The relation of density regulation to habitat specialization, evolution of a species' range, and the dynamics of biological invasions. *American Naturalist*, **172**, 233-247.

Fitzpatrick BM (2008) Dobzhansky-Muller model of hybrid dysfunction supported by poor burst-speed performance in hybrid tiger salamanders. *Journal of Evolutionary Biology*, **21**, 342-351.

Fitzpatrick LC (1976) Life history patterns of storage and utilization of lipids for energy in amphibians. *American Zoologist*, **16**, 725-732.

Flight PA, Nacci D, Champlin D, Whitehead A, Rand DM (2011) The effects of mitochondrial genotype on hypoxic survival and gene expression in a hybrid population of the killifish, *Fundulus heteroclitus*. *Molecular Ecology*, **20**, 4503-4520.

Garcia-Barros E, Romo Benito H (2010) The relationship between geographic range size and life history trait: is biogeographic history uncovered? A test using the Iberian butterflies. *Ecography*, **33**, 392-401.

Garcia-Ramos G, Kirkpatrick M (1997) Genetic models of rapid evolutionary divergence in peripheral populations. *Evolution*, **51**, 1-23.

Gaston KJ (2003) *The structure and dynamics of geographic ranges.* Oxford University Press, Oxford.

Gaston KJ (2009) Geographic range limits: achieving synthesis. *Proceedings of the Royal Society of London B*, **276**, 1395-1406.

Gaston KJ, Blackburn TM (1996) Conservation implications of geographic range size-body size relationships. *Conservation Biology*, **2**, 638-646.

Gause GF (1936) *The struggle for existence*. Williams and Wilkins, Baltimore.

Gavrilets S (1997) Hybrid zones with Dobzhansky-type epistatic selection. *Evolution*, **51**, 1027-1035.

Giordano A (2005) Effects of altitude on population structure and life history variation in the long-toed salamander. M.Sc. Thesis, Washington State University, Pullman.

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

Glor RE, Warren D (2010) Testing ecological explanations for biogeographic boundaries. *Evolution*, **65**, 673-683.

Godbout J, Fazekas A, Newton CH, Yeh FC, Bousquet J (2008) Glacial vicariance in the Pacific Northwest: evidence from a lodgepole pine mitochondrial DNA minisatellite for multiple genetically distinct and widely separated refugia. *Molecular Ecology*, **17**, 2463-2475.

Goebel AM, Ranker TA, Corn PS, Olmstead RG (2009) Mitochondrial DNA evolution in the *Anaxyrus boreas* species group. *Molecular Phylogenetics and Evolution*, **50**, 209-225.

Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology*, **19**, 3650-3663.

Goldberg EE, Lande R (2006) Ecological and reproductive character displacement on an environmental gradient. *Evolution*, **60**, 1344-1357.

Goldberg EE, Lande R. (2007) Species' borders and dispersal barriers. *American Naturalist*, **170,** 297-304.

Graham CH, Silva N, Velásquez-Tibatá J (2010) Evaluating the potential causes of range limits of birds of the Colombian Andes. *Journal of Biogeography*, **37**, 1863-1875.

Grant BR, Grant PR (1996) High survival of Darwin's finch hybrids: effects of beak morphology and diets. *Ecology*, **77**, 500-509.

Grayson KL, Cook LW, Todd MJ, Pierce D, Hopkins WA, Gatten RE Jr, Dorcas ME (2005) Effects of prey type on specific dynamic action, growth, and mass conversion efficiencies in the horned frog, *Ceratophrys cranwelli*. *Comparative Biochemistry and Physiology*, **141**, 298-304.

Haldane JBS (1956) The relation between density regulation and natural selection. *Proceedings of the Royal Society of London B*, **145**, 306-308.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series No.* **41**, 95-98.

Hamilton JA, Lexer C, Aitken SN (2013) Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* x *P. glauca*). *Molecular Ecology*, **22**, 827-841.

Harrison RG (1986) Pattern and process in a narrow hybrid zone. *Heredity*, **56**, 337-349.

Harrison RG (1990) Hybrid zones: windows on evolutionary process. *Oxford Surveys in Evolutionary Biology*, **7**, 69-128.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.

Hewitt G (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B*, **359**, 183-195.

Hewitt GM (1988) Hybrid zones—Natural laboratories for evolutionary studies. *Trends in Ecology and Evolution*, **3**, 158-167.

Hewitt GM (1989) The subdivision of species by hybrid zones. In: *Speciation and its Consequences* (eds Otte D, Endler JA), pp. 85-110. Sinauer Associates Inc, Sunderland, Massachusetts.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87-112.

Hijmans RJ, van Etten J (2012) raster: Geographic analysis and modeling with raster data. R package version 1.9-67.

Hijmans RJ, Phillips S, Leathwick J, Elith J (2012) dismo: Species distribution modeling. R package version 0.7-17.

Hochberg ME, Ives AR (1999) Can natural enemies enforce geographic range limits? *Ecography*, **22**, 268-276.

Hoffmann AA, Blows MW (1994) Species borders: ecological and evolutionary perspectives. *Trends in Ecology and Evolution*, **9**, 223-227.

Holt RD, Barfield M (2009) Trophic interactions and range limits: the diverse roles of predation. *Proceedings of the Royal Society of London B*, **276**, 1435-1442.

Holt RD, Keitt TH (2000) Alternative causes for range limits: a metapopulation perspective. *Ecology Letters*, **3**, 41-47.

Holt RD, Keitt TH (2005) Species' borders: a unifying theme in ecology. *Oikos*, **108**, 3-6.

Howard DJ (1986) A zone of overlap and hybridization between two ground cricket species. *Evolution*, **40**, 34-43.

Howard JH, Wallace RL (1985) Life history characteristics of populations of the Long-toed salamander (*Ambystoma macrodactylum*) from different altitudes. *American Midland Naturalist*, **113**, 361-373.

Howard JH, Wallace RL, Stauffer JR Jr (1983) Critical thermal maxima in populations of *Ambystoma macrodactylum* from different elevations. *Journal of Herpetology*, **17**, 400-402.

Hua X, Wiens JJ (2013) How does climate influence speciation? *American Naturalist*, **182**, 1-12.

Hutchinson GE (1957) Concluding remarks. *Cold Spring Harbour Symposia on Quantitative Biology*, **22**, 415-427.

Imsland AK, Foss A, Stefansson SO (2001) Variation in food intake, food conversion efficiency and growth of juvenile turbot from different geographic strains. *Journal of Fish Biology*, **59**, 449-454.

Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution*, **56**, 2383-2394.

Irwin DE (2012) Local adaptation along smooth ecological gradients causes phylogeographic breaks and phenotypic clustering. *American Naturalist*, **180**, 35-49.

Irwin DE, Alström P, Olsson U, Benowitz-Fredericks ZM (2001) Cryptic species in the genus *Phylloscopus* (Old World leaf warblers). *Ibis*, **143**, 233-247.

Irwin DE, Bensch S, Price TD (2001) Speciation in a ring. *Nature*, **409**, 333-337.

Irwin DE, Rubtsov AS, Panov EN (2009) Mitochondrial introgression and replacement between yellowhammers (*Emberiza citrinella*) and pine buntings (*E leucocephalos*; Aves Passeriformes). *The Biological Journal of the Linnean Society*, **98**, 422-438.

IUCN, Conservation International & Natureserve, 2008. IUCN Red List of Threatened Species: Global amphibian assessment. Version 2008.1. http://www.iucnredlist.org/amphibians/

Jackson ST, Overpeck JT (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, **26**, 194-220.

Jankowski JE, Robinson SK, Levey DJ (2010) Squeezed at the top: interspecific aggression may constrain elevational ranges in tropical birds. *Ecology*, **91**, 1877-1884.

Johnstone JF, Chapin FS (2003) Non-equilibrium succession dynamics indicate continued northern migration of lodgepole pine. *Global Change Biology*, **9**, 1401-1409.

Jonassen J M, Imsland AK, Fitzgerald R, Bonga SW, Ham EV, Naevdal G, Stefánsson MO, Stefansson SO (2000) *Journal of Fish Biol*ogy, **56**, 279-294.

Kaplan RH, Salthe SN (1979) The allometry of reproduction: an empirical view in salamanders. *American Naturalist,* **113**, 671-689.

Kearney M, Porter, W (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334-350.

Keitt TH, Lewis MA, Holt RD (2001) Allee effects, invasion pinning, and species' borders. *American Naturalist*, **157**, 203-216.

Keller I, Seehausen O (2012) Thermal adaptation and ecological speciation. *Molecular Ecology*, **21**, 782-799.

Key KHL (1968) The concept of stasipatric speciation. *Systematic Zoology*, **17**, 14-22.

Key KHL (1982) Species, parapatry, and the morabine grasshoppers. *Systematic Zoology*, **30**, 425-458.

Kezer J, Farner DS (1955) Life history patterns of the salamander *Ambystoma macrodactylum* in the high Cascade mountains of southern Oregon. *Copeia*, **1955**, 127-131.

Kharouba HM, McCune JL, Thuiller W, Huntley B (2013) Do ecological differences between taxonomic groups influence the relationship between species' distributions and climate? A global meta-analysis using species distribution models. *Ecography*, **36**, 657-664.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *American Naturalist*, **150**, 1-23.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623-2635.

Kozak KH, Wiens JA (2006) Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution*, **60**, 2604-2621.

Krosby M, Rohwer S (2009) A 2000 km genetic wake yields evidence for northern glacial refugia and hybrid zone movement in a pair of songbirds. *Proceedings of the Royal Society of London B*, **276**, 615-621.

Kruuk LEB, Gilchrist JS, Barton NH (1999) Hybrid dysfunction in Fire-Bellied Toads (Bombina). *Evolution*, **53**, 1611-1616.

Kuchta SR, Tan A (2005) Isolation by distance and post-glacial range expansion in the rough-skinned newt *Taricha granulose*. *Molecular Ecology*, **14**, 225-244.

Latch EK, Kierepka EM, Heffelfinger JR, Rhodes OE Jr. (2011) Hybrid swarm between diverent lineages of mule deer (*Odocoileus hemionus*). *Molecular Ecology*, **20**, 5265-5279.

Lee-Yaw JA, Irwin DE (2012). Large geographic range size reflects a patchword of divergent lineages in the long-toed salamander(*Ambystoma macrodactylum*). *Journal of Evolutionary Biology*, **25**, 2276-2287.

Lee-Yaw JA, Irwin JT, Green DM (2008) Postglacial range expansion from northern refugia by the wood frog *Rana sylvatica*. *Molecular Ecology,* **17**, 867-884.

Leger EA, Forister ML (2009) Colonization abundance and geographic range size of gravestone lichens. *Basic and Applied Ecology*, **10**, 279-287.

Lemmon EM, Lemmon AR (2010) Reinforcement in chorus frogs: lifetime fitness estimates including intrinsic natural selection and sexual selection against hybrids. *Evolution*, **64**, 1748-1761.

Lester SE, Ruttenberg BI, Gaines SD, Kinlan BP (2007) The relationship between dispersal ability and geographic range size. *Ecology Letters*, **10**, 745-758.

Letcher AJ, Purvis A, Nee S, Harvey PH (1994) Patterns of overlap in the geographic ranges of Palearctic and British mammals. *The Journal of Animal Ecology*, **63**, 871-879.

Levsen ND, Tiffin P, Olson MS (2012) Pleistocene speciation in the genus *Populus* (Salicaceae). *Systematic Biology*, **61**, 401-412.

Li J, Heap AD (2008) A review of spatial interpolation methods for environmental scientists. Geoscience Australia Records 2008/23.

Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, **11**, 995-1007.

Loxdale HD, Lushai G, Harvey JA (2011) The evolutionary improbability of 'generalism' in nature with special reference to insects. *The Biological Journal of the Linnean Society*, **103**, 1-18.

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91-99.

MacArthur RH (1972) *Geographical Ecology*. Harper and Row, New York.

Makowsky R, Chesser J, Rissler LJ (2008) A striking lack of genetic diversity across the wide-ranging amphibian *Gastrophryne carolinensis* (Anura: Microhylidae). *Genetica*, **135**, 169-183.

Masta SE, Sullivan BK, Lamb T, Routman EJ (2002) Molecular systematics hybridization and phylogeography of the *Bufo americanus* complex in Eastern North America. *Molecular Phylogenetics and Evolution*, **24**, 302-314.

Mathis A (1990) Territoriality in a terrestrial salamander: the influence of resource quality and body size. *Behaviour*, **112**, 162-175.

Matsui M, Shimada T, Lui W, Maryati M, Khonsue W, Orlov N (2006) Phylogenetic relationships of Oriental torrent frogs in the genus *Amolops* and its allies (Amphibia Anura Ranidae). *Molecular Phylogenetics and Evolution*, **38**, 659-666.

Maynard Smith J, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genetics Research*, **23**, 23-35.

Mayr E (1963) *Animal Species and Evolution*. Harvard University Press, Cambridge MA.

McCormack JE, Zellmer AJ, Knowles LL (2009) Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation?: Insights from tests with niche models. *Evolution*, **64**, 1231-1244.

Mila B, Carranza S, Guillaume O, Clobert J (2010) Marked genetic structuring and extreme dispersal limitation in the Pyrenean brook newt *Calotriton asper* (Amphibia: Salamandridae) revealed by genome-wide AFLP but not mtDNA. *Molecular Ecology*, **19**, 108-120.

Mittleman MB (1948) American Caudata II Geographic variation in *Ambystoma macrodactylum*. *Herpetologica*, **4**, 81-95.

Moore WS (1977) An evaluation of narrow hybrid zones in vertebrates. *Quarterly Review of Biology*, **52**, 263-277.

Mueller RI, Macey JR, Jaekel M, Wake DB, Boore JL (2004) Morphological homoplasy life history evolution and historical biogeography of plethodontid salamanders inferred from

complete mitochondrial genomes. *Proceedings of the National Academy of Sciences of the USA*, **101**, 13820-13825.

Munguía M, Rahbek C, Rangel TF, Diniz-Filho JAF, Araújo MB (2012) Equilibrium of global amphibian species distributions with climate. PLoS ONE 7, e34420. doi:10.1371/journal.pone.0034420.

Murphy MA, Evans JS, Cushman SA, Storfer A (2008) Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. *Ecography*, **31**, 685-697.

Niedzialkowska J, Jedrzejewska B, Honnen A, Otto T, Sideorovich VE, Perzanowski K *et al* (2011) Molecular biogeography of red deer *Cervus elaphus* from eastern Europe: insights from mitochondrial DNA sequences. *Acta Theriologic,* **56**, 1-12.

Nielson M, Lohman K, Daugherty CH, Allendorf FW, Knudsen KL, Sullivan J (2006) Allozyme and mitochondrial DNA variation in the tailed frog (Anura: *Ascaphus*): The influence of geography and gene flow. *Herpetologica*, **62**, 235-258.

Nolte AW, Gompert Z, Buerkle CA (2009) Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Molecular Ecology*, **18**, 2615-2627.

Nürnberger B, Barton N, MacCallum C, Gilchrist J, Appelby M (1995) Natural selection on quantitative traits in the *Bombina* hybrid zone. *Evolution*, **49**, 1224-1238.

Oksanen J, Blanchet GF, Kindt R, Legendre P, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2011) vegan: Community Ecology Package. R package version 1.17-9.

Oliver PM, Adams M, Lee MSY, Hutchinson MN, Doughty P (2009) Cryptic diversity in vertebrates: molecular data double estimates of species diversity in a radiation of Australian lizards (*Diplodactylus* Gekkota). *Proceedings of the Royal Society of London B*, **276**, 2001-2007.

Paradis E, Claude J, Strimmer K (2004) APE: analysis of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289-290.

Parris MJ (2001) Hybridization in leopard frogs (Rana pipiens complex): variation in interspecific hybrid larval fitness components along a natural contact zone. *Evolutionary Ecology Research*, **3**, 91-105.

Paul JR, Morton C, Taylor CM, Tonsor SJ (2009) Evolutionary time for dispersal limits the extent but not the occupancy of species' potential ranges in the tropical plant genus *Psychotria* (Rubiaceae). *American Naturalist*, **173**, 188-199.

Peterson, AT (2011) Ecological niche conservatism: a time-structured review of evidence. Journal of Biogeography, **38**, 817-827.

Peterson AT, Soberón J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265-1267.

Peterson AT, Soberón J, Pearson RG, Anderson RP, Marínez-Meyer E, Nakamura M, Araújo MB (2011) *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton.

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology and Evolution*, **24**, 386-393.

Pfennig DW, Murphy PJ (2000) Character displacement in polyphonic tadpoles. *Evolution*, **54**, 1738-1749.

Pfennig KS, Pfennig DW (2005) Character displacement as the "best of a bad situation": fitness trade-offs resulting from selection to minimize resource and mate competition. *Evolution*, **59**, 2200-2208.

Pfennig KS, Pfennig DW (2009) Character displacement: ecological and reproductive responses to a common evolutionary problem. *Quarterly Review of Biology*, **84**, 253-276.

Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evolutionary Biology*, **7**, 121.

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161-175.

Phillips SJ (2008) Transferability, sample seletion bias and background data in presence-only modelling: a response to Peterson *et al.* (2007). *Ecography*, **31**, 272-278.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modeling*, **190**, 231-259.

Pigot AL, Tobias JA (2013) Species interactions constrain geographic range expansion over evolutionary time. *Ecology Letters*, **16**, 330-338.

Pinheiro J, Bates D, DebRoy S, Sarkar D, the R Development Core Team (2011) nlme: Linear and Nonlinear mixed effects models. R package version 3.1-100.

Pollard KS, Gilbert HN, Ge Y, Taylor S, Dudoit S (2005). multtest: Resampling-based multiple hypothesis testing. R package version 2.8.0.

Price TD, Kirkpatrick M (2009) Evolutionary stable range limits set by interspecific competition. *Proceedings of the Royal Society of London B*, **276**, 1429-1434.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Purchase CF, Brown JA (2000) Interpopulation differences in growth rates and food conversion efficiencies of young Grand Banks and Gulf of Maine Atlantic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences*, **57**, 2223-2229.

Putta S, Smith JJ, Walker JA, Rondet M, Weisrock DW, Monaghan J, Samuels AK, Kump K, King DC, Maness NJ, Habermann B, Tanaka E, Bryant SV, Gardiner DM, Parichy DM, Voss SR (2004) From biomedicine to natural history research: EST resources for ambystomid salamanders. *BMC Genomics*, **5**, 54.

Raes N, ter Steege H (2007) A null-model for significance testing of presence-only species distribution models. *Ecography*, **30**, 727-736.

Rauch EM, Bar-Yam Y (2004) Theory predicts the uneven distribution of genetic diversity within species. *Nature*, **431**, 449-452.

Ricklefs RE (2010) Host-pathogen coevolution, secondary sympatry and species diversification. *Philosophical Transactions of the Royal Society of London B*, **65**, 1139-1147.

Rissler LJ, Apodaca JJ (2007) Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the Black salamander (*Aneides flavipunctatus*). *Systematic Biology*, **56**, 924-942.

Rissler LJ, Smith WH (2010) Mapping amphibian contact zones and phylogeographical break hotspots across the United States. *Molecular Ecology*, **19**, 5404-5416.

Sagarin RD, Gaines SD (2002) The 'abundant' center' distribution: to what extent is it a biogeographical rule? *Ecology Letter*, **5**, 137-147.

Sage RD, Heyneman D, Lim K, Wilson AC (1986) Wormy mice in a hybrid zone. *Nature*, **324**, 60-63.

Samis KE, Eckert CG (2009) Ecological correlates of fitness across the northern geographic range limit of a Pacific Coast dune plant. *Ecology*, **90**, 3051-3061.

Savage WK (2008) Landscape genetics and phylogenetic relationships in the *Ambystoma macrodactylum* species complex. PhD Dissertation, The University of California, Davis.

Savage WK, Fremier AK, Shaffer HB (2010) Landscape genetics of alpine Sierra Nevada salamanders reveal extreme population subdivision in space and time. *Molecular Ecology*, **19**, 3301-3314.

Schluter D (1995) Adaptive radiation in sticklebacks: trade-offs in feeding performance and growth. *Ecology*, **76**, 82-90.

Schluter D (2009) Evidence for ecological speciation and its alternative. *Science*, **323**, 737-741.

Schnitzler J, Graham CH, Dormann CF, Schiffers K, Linder HP (2012) Climatic niche evolution and species diversification in the Cape flora, South Africa. *Journal of Biogeography*, **39**, 2201-2211.

Scott DE, Fore MR (1995) The effect of food limitation on lipid levels, growth, and reproduction in the marbled salamander, *Ambystoma opacum*. *Herpetologica*, **51**, 462-471.

Servedio MR, Noor MAF (2003) The role of reinforcement in speciation: Theory and data. *Annual Review* of *Ecology* and Systematics, **34**, 339-364.

Sexton JP, Strauss SY, Rice KJ (2011) Gene flow increases fitness at the warm edge of a species' range. *Proceedings of the National Academy of Sciences*, **108**, 11704-11709.

Sexton JP, McIntyre PJ, Angert AL, Rice KJ (2009) Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution and Systematics*, **40**, 415-436.

Shafer ABA, Cullingham CI, Cote SD, Coltman DW (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**, 4589-4621.

Singhal S, Moritz C (2012) Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard. *Evolution*, **66**, 1474-1489.

Smith TB, Skulason S (1996) Evolutionary significance of resource polymorphisms in fishes, amphibians, and birds. *Annual Review of Ecology and Systematics*, **27**, 111-133.

Soltis DE, Gizendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution,* **206**, 353-373.

Soto-Centeno JA, Barrow LN, Allen JM, Reed DL (2013) Reevaluation of a classic phylogeographic barrier: new techniques reveal the influence of microgeographic climate variation on population divergence. *Ecology and Evolution*, **3**, 1603-1613.

Stanton-Geddes J, Tiffin P, Shaw RG (2012) Role of climate and competitors in limiting fitness across range edges of an annual plant. *Ecology*, **93**, 1604-1613.

Steele CA, Carstens BC, Storfer A, Sullivan J (2005) Testing hypotheses of speciation timing in *Dicamptodon copei* and *Dicamptodon aterriumus* (Caudata: Dicamptodonitdae). *Molecular Phylogenetics and Evolution*, **36**, 90-100.

Steele CA, Storfer A (2006) Coalescent-based hypothesis testing supports multiple Pleistocene refugia in the Pacific Northwest of the Pacific giant salamander (*Dicamptodon tenebrosus*). *Molecular Ecology*, **15**, 2477-2487.

Stevens VM, Coulon A (2012) Landscape effects on spatial dynamics: the natterjack toad as a case study. In: *Dispersal Ecology and Evolution* (eds Clobert J, Baguette M, Benton TG, Bullock JM), pp. 280-289. Oxford University Press, Oxford.

Svenning JC, Skov F (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, **7**, 565-573.

Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *American Naturalist*, **166**, 581-591.

Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285-1293.

Takahashi MK, Parris MJ (2008) Life cycle polyphenism as a factor affecting ecological divergence within *Notophthalmus viridescens*. *Oecologia*, **158**, 23-34.

Tallmon DA, Funk WC, Dunlap WW, Allendorf FW (2000) Genetic differentiation among long-toed salamander (*Ambystoma macrodactylum*) populations. *Copeia*, **1**, 27-35.

Tallmon DA, Funk WC, Dunlap WW, Allendorf FW (2000) Genetic differentiation among long-toed salamander (*Ambystoma macrodactylum*) populations. *Copeia*, **1**, 27-35.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512-526.

Tan DSH, Ang Y, Lim GS, Ismail MRB, Meier R (2010) From 'cryptic species' to integrative taxonomy: an iterative process involving DNA sequences, morphology and behaviour leads to the resurrection of *Sepsis pyrrhosoma* (Sepsidae: Diptera). *Zoologica Scripta*, **39**, 51-61.

Taylor EB, Gerlinsky C, Farrell N, Gow JL (2012) A test of hybrid growth disadvantage in wild, free-ranging species pairs of threespine stickleback (*Gasterosteus aculeatus*) and its implications for ecological speciation. *Evolution*, **66**, 240-251.

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680.

Thompson MD (2003) Phylogeography of the Long-Toed Salamander *Ambystoma macrodactylum*. MSc Dissertation, The University of Calgary.

Thompson M, Russel A (2005) Glacial Retreat and its Influence on Migration of Mitochondrial Genes in the Long-toed Salamander (*Ambystoma macrodactylum*) in Western North America. In: *Migration of Organisms* (ed. ElewaAMT), pp. 205–246. Springer, Berlin Heidelberg.

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology,* **21**, 3907-3930.

Toews DPL, Irwin DE (2008) Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. *Molecular Ecology*, **17**, 2691-2705.

Turner LM, Schwahn DJ, Harr B (2011) Reduced male fertility is common but highly variable in form and severity in a natural house mouse hybrid zone. *Evolution*, **66**, 443-458.

VanDerWal, J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological Modelling,* **220**, 589-594.

Vekemans X (2002) AFLP-SURV version 1.0. Laboratoire de Genetique et Ecologie Vegetale, Universite Libre de Bruxelles, Belgium. http://www.ulb.ac.be/sciences/lagev/aflp-surv.html

Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139-151.

Verrell PA (2006) The female reproductive cycle of the North American salamander *Ambystoma macrodactylum columbianum*. *Amphibia-Reptilia*, **27**, 274-277.

Vincze T, Posfai J, Roberts RJ (2003). NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Research*, **31**, 3688-3691.

Vos P, Hogers R, Bleeker M, Reijans M van de, Lee T, Hornes M *et al* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407-4414.

Voss SR, Shaffer HB (1997) Adaptive evolution via a major gene effect: Paedomorphosis in the Mexican axolotl. *Proceedings of the National Academy of Sciences of the USA*, **94**, 14185-14189.

Voss SR, Smith JJ, Gardiner DM, Parichy DM (2001) Conserved vertebrate chromosome segments in the large salamander genome. *Genetics*, **158**, 735-746.

Wagner RS, Miller MP, Crisafulli CM, Haig SM (2005) Geographic variation, genetic structure and conservation unit designation in the Larch Mountain salamander (*Plethodon larselli*). *Canadian Journal of Zoology*, **83**, 396-406.

Waldron A (2010) Lineages that cheat death: surviving the squeeze on range size. *Evolution*, **64**, 2278-2292.

Wang T, Hemann A, Spittlehouse D, Murdock TN (2012) ClimateWNA – High-resolution spatial climate data for Western North America. *Journal of Applied Meteorology and Climatology*, **61**, 16-29.

Weir JT, Schluter D (2004) Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society of London B*, **271**, 1881-1887.

Weisrock DW, Shaffer HB, Storz BL, Storz SR, Voss SR (2006) Multiple nuclear gene sequences identify phylogenetic species boundaries in the rapidly radiating clade of Mexican ambystomid salamanders. *Molecular Ecology*, **15**, 2489-2503.

Werner P, Lötters S, Schmidt BR, Engler JO and Rödder D. (2013) The role of climate for the range limits of parapatric European land salamanders. *Ecography*, **36**, 1-11.

Wheeler B (2010) lmPerm: Permutation tests for linear models. R package version 1.1-2.

Wiens JJ (2004) Speciation and ecology revisted: phylogenetic niche conservatism and the origins of species. *Evolution*, **58**, 193-197.

Wiens JJ (2011) The niche, biogeography and species interactions. *Philosophical Transactions of the Royal Society of London B,* **366**, 2336-2350.

Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Davies J, Grytnes JA, Harrison SP, Hawkins BA, Holt RD, McCain CM, Stephens PR (2010) Niche conservatism as an emerging principle in ecology and evolution. *Ecology Letters*, **13**, 1310-1324.

Wiens JJ, Engstrom TN, Chippindale PT (2006) Rapid diversification, incomplete isolation, and the "speciation clock" in North American salamanders (Genus *Plethodon*): Testing the hybrid swarm hypothesis of rapid radiation. *Evolution*, **60**, 2585-2603.

Wiens JJ, Kuczynski CA, Stephens PR (2010) Discordant mitochondrial and nuclear gene phylogenies in emydid turtles: implications for speciation and conservation. *The Biological Journal of the Linnean Society*, **99**, 445-461.

Wildy EL, Chivers DP, Kiesecker JM, Blaustein AR (2001) The effects of food level and conspecific density on biting and cannibalism in larval long-toed salamanders, *Ambystoma macrodactylum*. *Oecologia*, **128**, 202-209.

Wirtz P (1999) Mother species-father species: unidirectional hybridization in animals with female choice. *Animal Behaviour*, **58**, 1-12.

Wu CI (2001) The genic view of the process of speciation. *Evolutionary Biology*, **14**, 851-865.

Zamudio KR, Savage WK (2003) Historical isolation range expansion and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution*, **57**, 1631-1652.

APPENDICES

Appendix A Supplementary material for Chapter 2

A.1 Locality data for populations surveyed

Table A1. Locality information for samples used in genetic survey of *cyt b* haplotypes and AFLP data (GenBank Accession numbers: JX650148-JX650223). N_{mtDNA} refers to number of samples sequenced; all AFLP samples (N_{AFLP}) were either sequenced or assigned to mtDNA group based on PCR-RFLP. STRUCTURE plot order lists the order in which populations are presented in Figure 2.2.

Collector_Site	Locality Description	County	State	Num (AFLP)	Num (mtDNA)	Haplotypes	STRUCTURE Plot Order
JLee-Yaw_1	Wilson Wildlife Area	Benton	OR	1	0	N/A	15
JLee-Yaw_2	Tangent Rd	Linn	OR	3	0	N/A	16
JLee-Yaw_3	Bond Road near town of Lebanon	Linn	OR	1	0	N/A	19
JLee-Yaw_4	Evergreen Campus	Thurston	WA	2	1	hap1(1)	7
JLee-Yaw_6	Smoot Hill	Whitman	WA	5	1	hap3(1)	61
JLee-Yaw_8	Hansen Elementary School in West Olympia	Thurston	WA	1	0	N/A	8
JLee-Yaw_9	Gifford-Pinchot National Forest	Lewis	WA	3	0	N/A	11
JLee-Yaw_10	Eugene	Lane	OR	2	1	hap2(1)	17
JLee-Yaw_11	Virgil Phillips Farm	Latah	ID	10	2	hap3(2)	62
JLee-Yaw_12	Fairview Wetlands	Marion	OR	2	1	hap4(1)	18
JLee-Yaw_13	Du Pont Village	Pierce	WA	2	1	hap5(1)	9

Collector_Site	Locality Description	County	State	Num (AFLP)	Num (mtDNA)	Haplotypes	STRUCTURE Plot Order
JLee-Yaw_14	Wilkeson	Pierce	WA	5	1	hap6(1)	10
JLee-Yaw_15	Ellensburg	Kittitas	WA	9	1	hap7(1)	14
						hap8(1).	58
JLee-Yaw_16	Quincy Lakes	Grant	WA	2	2	Hap9(1)	
JLee-Yaw_17	Fishtrap 1	Spokane	WA	2	0	N/A	59
						hap10(1),	60
						hap11(2),	
JLee-Yaw_18	Fishtrap 2	Lincoln	WA	3	4	hap12(1)	
JLee-Yaw_19	Gun Club Road	Latah	ID	1	1	hap3(1)	63
						hap13(1),	64
JLee-Yaw_20	Bovil	Latah	ID	9	2	hap14(1)	
JLee-Yaw_21	Midvale Walking Trail	Washington	ID	7	0	N/A	108
JLee-Yaw_23	outside Baker City	Baker	OR	7	1	hap15(1)	107
JLee-Yaw_24	Malheur National Forest	Grant	OR	5	1	hap16(1)	106
JLee-Yaw_25	Leasy Rd outside Long Creek	Grant	OR	2	0	N/A	104
JLee-Yaw_26	East of Mt Vernon	Grant	OR	1	0	N/A	105
JLee-Yaw_27	Juniper Canyon Rd-Prineville	Crook	OR	8	1	hap17(1)	103
JLee-Yaw_28	Sunriver	Deschutes	OR	4	1	hap18(1)	22
JLee-Yaw_29	Indian Ford Rd outside Sisters	Deschutes	OR	2	0	N/A	20
JLee-Yaw_30	NE side of Black Butte	Jefferson	OR	3	0	N/A	21
						hap19(1),	13
JLee-Yaw_31	North of Ellensburg	Kittitas	WA	10	2	hap20(1)	

Collector_Site	Locality Description	County	State	Num (AFLP)	Num (mtDNA)	Haplotypes	STRUCTURE Plot Order
	Westside Rd and Woods-Steele						12
JLee-Yaw_32	Rd south of S. Cle Elum	Kittitas	WA	5	0	N/A	
						hap20(4),	6
JLee-Yaw_34	Keremeos	Okanagan Region	BC	5	5	hap21(1)	
JLee-Yaw_35	Pole Cutter Creek	Okanagan Region	BC	7	5	hap20(5)	5
						hap20 (1),	4
JLee-Yaw_36	Harrison Hot Springs	Lower Mainland Region	BC	1	2	hap22(1)	
JLee-Yaw_38	10 mile Lake Park	Cariboo Region	BC	4	1	hap3(1)	29
JLee-Yaw_39	Forest for the World UNBC	Omineca Region	BC	4	1	hap3(1)	30
	Hwy 16_ ~6.2 km east of Prince						31
JLee-Yaw_41	George	Omineca Region	BC	2	0	N/A	
JLee-Yaw_42	Hwy 16 across from rest area	Omineca Region	BC	1	0	N/A	32
JLee-Yaw_43	McBride	Omineca Region	BC	2	1	hap23(1)	33
JLee-Yaw_44	Hwy 5 ~14 km south of Hwy 16	Omineca Region	BC	1	0	N/A	34
JLee-Yaw_45	Jasper/Hinton Airport	Yellowhead	AB	2	1	hap3(1)	35
JLee-Yaw_46	Frank	Ranchlands	AB	0	3	hap24(3)	N/A
						hap24(1),	71
						hap25(3),	
JLee-Yaw_47	Sparwood	Kootenay Region	BC	6	5	hap26(1)	
						hap27(1),	55
						hap28(1),	
JLee-Yaw_48	Cranbrook	Kootenay Region	BC	10	5	hap29(3)	

Collector_Site	Locality Description	County	State	Num (AFLP)	Num (mtDNA)	Haplotypes	STRUCTURE Plot Order
						hap29(2),	54
						hap30(1),	
						hap31(2),	
JLee-Yaw_49	Yahk	Kootenay Region	BC	9	6	hap32(1)	
						hap33(1),	53
						hap34(1),	
						hap35(1),	
JLee-Yaw_50	Creston	Kootenay Region	BC	4	4	hap36(1)	
	Creston Valley Wildlife						51
JLee-Yaw_51	Management Area	Kootenay Region	BC	2	0	N/A	
						hap34(2),	52
	Creston Valley Wildlife					hap37(1),	
JLee-Yaw_52	Management Area	Kootenay Region	BC	4	4	hap38(1)	
						hap39(1),	50
						hap40(1),	
						hap41(1),	
						hap42(1),	
JLee-Yaw_53	Salmo	Kootenay Region	BC	9	5	hap43(1)	
						hap29(2),	49
						hap31(1),	
						hap44(1),	
JLee-Yaw_54	Castlegar	Kootenay Region	BC	8	5	hap45(1)	
JLee-Yaw_55	Greenwood	Kootenay Region	BC	8	4	hap3(4)	48

Collector_Site	Locality Description	County	State	Num (AFLP)	Num (mtDNA)	Haplotypes	STRUCTURE Plot Order
JLee-Yaw_56	Beaver Creek Divide	Clearwater	ID	7	2	hap46(2)	76
JLee-Yaw_57	Whitehouse Pond	Idaho	ID	3	1	hap47(1)	77
JLee-Yaw_58	Glen Lake	Ravalli	MT	2	1	hap48(1)	90
JLee-Yaw_59	Glen Lake_Second Pond	Ravalli	MT	4	2	hap48(1), hap49(1)	91
JLee-Yaw_60	Lolo Forest	Missoula	MT	2	3	hap24(2), hap50(1)	93
JLee-Yaw_61	Ogden Mt Rd	Powell	MT	3	3	hap24(3)	94
JLee-Yaw_62	Aspen Grove Campground	Lewis and Clark	MT	4	1	hap24(1)	95
JLee-Yaw_64	Mosquito Flat Reservoir	Custer	ID	5	3	hap51(2), hap52(1)	101
JLee-Yaw_65	Cape Horn Lakes	Custer	ID	3	0	N/A	102
JLee-Yaw_66	Bull Trout Lake Rd	Boise	ID	5	1	hap53(1)	97
JLee-Yaw_67	Skunk Creek Rd	Valley	ID	4	1	hap54(1)	96
JLee-Yaw_68	Whistler Athletes Village	Lower Mainland Region	BC	2	0	N/A	1
JLee-Yaw_69	Cypress Hill Hollybrun Mt	Lower Mainland Region	BC	1	5	hap20(3), hap55(1), hap56(1),	3
JLee-Yaw_93	Mabel Lake ditch 1	Okanagan Region	BC	2	0	N/A	40
JLee-Yaw_94	Mabel Lake ditch 2	Okanagan Region	BC	2	0	N/A	42
JLee-Yaw_95	Sugar Lake ditch	Okanagan Region	BC	2	0	N/A	43
JLee-Yaw_105	Akamina Pass	Waterton	AB	4	1	hap57(1)	80
JLee-Yaw_119	Birds Eye Butte Rd	Waterton	AB	2	0	N/A	82

Collector_Site	Locality Description	County	State	Num (AFLP)	Num (mtDNA)	Haplotypes	STRUCTURE Plot Order
JLee-Yaw_127	Castle Cliff Pond	Banff	AB	1	0	N/A	67
JLee-Yaw_128	Hector Lake	Banff	AB	2	1	hap58(1)	66
JLee-Yaw_129	Hwy 1A towards Exsaw	Banff	AB	5	0	N/A	69
JLee-Yaw_145	Geraldine Lakes Trail	Jasper	AB	5	1	hap59(1)	36
JLee-Yaw_151	Chancellor Campground	Kootenay Region	BC	8	1	hap58(1)	47
JLee-Yaw_152	Beaver River Trail Rd	Kootenay Region	BC	10	1	hap60(1)	46
JLee-Yaw_153	Skunk Cabbage boardwalk	Kootenay Region	BC	11	0	N/A	45
JLee-Yaw_154	Mizon Rd	Thompson Region	BC	9	0	N/A	41
JLee-Yaw_155	Road to Donald Lake	Thompson Region	BC	5	0	N/A	23
JLee-Yaw_156	Road to Jimmy Lake	Thompson Region	BC	5	0	N/A	24
JLee-Yaw_157	Roosevelt	Klickitat	WA	0	2	hap61(2)	N/A
AGiordano_PRT	N/A	Lincoln	MT	1	0	N/A	74
AGiordano_SL	N/A	Lincoln	MT	2	0	N/A	85
AGiordano_410	N/A	Lincoln	MT	2	0	N/A	56
AGiordano_BAM	N/A	Lincoln	MT	2	0	N/A	83
AGiordano_Bear	N/A	Lincoln	MT	1	0	N/A	86
AGiordano_BLH	N/A	Lincoln	MT	2	0	N/A	72
AGiordano_BOHS	N/A	Lincoln	MT	3	0	N/A	73
AGiordano_TBP	N/A	Lincoln	MT	1	0	N/A	57
AGiordano_LEL	N/A	Lincoln	MT	2	0	N/A	75
AGiordano_MEL	N/A	Lincoln	MT	2	0	N/A	84
DPilliod_StriderLake	N/A	Lemhi	ID	1	1	hap62(1)	100
DPilliod_GreggsLake	N/A	Lemhi	ID	1	0	N/A	99

Collector_Site	Locality Description	County	State	Num (AFLP)	Num (mtDNA)	Haplotypes	STRUCTURE Plot Order
DPilliod_UpperCache	N/A	Lemhi	ID	1	0	N/A	98
Jweir_SantaCruz	N/A	Santa Cruz	CA	0	1	hap74(1)	N/A
						hap3(1),	39
MThompson_MT116	N/A	Thompson Region	BC	1	2	hap69(1)	
						hap3(1),	N/A
MThompson_MT118	N/A	Thompson Region	BC	0	2	hap20(1)	
						hap20(2),	N/A
MThompson_MT117	N/A	Thompson Region	BC	0	3	hap70(1)	
MThompson_MT20	N/A	Flathead	MT	0	2	hap24(2)	N/A
MThompson_MT41	N/A	Banff	AB	0	1	hap24(1)	N/A
						hap63(2),	N/A
MThompson_MT4	N/A	Sanders	MT	0	3	hap64(1)	
MThompson_MT81	N/A	Kootenay Region	BC	1	1	hap63(1)	65
MThompson_MT84	N/A	Kootenay Region	BC	0	3	hap63(3)	N/A
MThompson_MT62	N/A	Omineca Region	BC	0	1	hap65(1)	N/A
						hap66(1),	28
MThompson_MT51	N/A	Fairview	AB	1	2	hap67(1)	
MThompson_MT104	N/A	Skeena Region	BC	0	1	hap68(1)	N/A
MThompson_MT121	N/A	North Okanagon	BC	2	3	hap71(3)	44
						hap3(1),	N/A
						hap72(1),	
MThompson_MT122	N/A	Thompson Region	BC	0	3	hap73(1)	
MThompson_MT100	N/A	Skeena Region	BC	2	0	N/A	25

Collector_Site	Locality Description	County	State	Num (AFLP)	Num (mtDNA)	Haplotypes	STRUCTURE Plot Order
MThompson_MT106	N/A	Omenica Region	BC	1	0	N/A	26
MThompson_MT109	N/A	Caribou Region	BC	1	0	N/A	37
MThompson_MT111	N/A	Caribou Region	BC	1	0	N/A	38
MThompson_MT112	N/A	Thompson Region	BC	2	0	N/A	2
MThompson_MT16	N/A	Flathead	MT	3	0	N/A	88
MThompson_MT19	N/A	Glacier	MT	1	0	N/A	89
MThompson_MT21	N/A	Flathead	MT	1	0	N/A	87
MThompson_MT26	N/A	Waterton	AB	1	0	N/A	81
MThompson_MT31	N/A	Pincher Creek	AB	1	0	N/A	79
MThompson_MT35	N/A	Banff	AB	5	0	N/A	70
MThompson_MT5	N/A	Missoula	MT	2	0	N/A	92
MThompson_MT50	N/A	Fairview	AB	1	0	N/A	27
MThompson_MT87	N/A	Kootenay Region	BC	4	0	N/A	68
MThompson_MT93	N/A	Kootenay Region	BC	2	0	N/A	78
MVZ237485	E Port Angeles	Clallam	WA	0	1	hap75(1)	N/A
MVZ235936	NE Ebbett's Pass	Alpine	CA	0	1	hap76(1)	N/A

A.2 PCR-RFLP protocol for diagnosing different mtDNA lineages

Step 1: PCR

25 µl reactions consisting of:

0.25 µl dNTPs (10 mM)
2.5 µl 10x Buffer (NEB)
0.75 µl mgCl (50 mM)
1 µl R primer (10 µM) Amb_pleth_cytb_R (5' YCR RTT TTC GRC TTA CAA GG)
1 µl F primer (10 µM) Amac_rflp_cytb (5' GAC AAA GCT ACY TTA ACT CG)
0.1µl TAQ
18.4 µl ddH2O
1.0 µl DNA

PCR program:

94°C for 2 mins
35 cycles of:
{94°C for 30 s
56°C for 45 s
72°C for 1 min 10 s}
72°C for 7 mins

Step 2: EcoRI digest

6 µl reactions consisting of:

0.4 µl EcoRI buffer (NEB)
0.033 µl EcoRI (NEB: R0101S)
2 µl PCR product from step 1
3.57 µl ddH2O

Incubation:

37°C for 2 hours

Interpretation of Results:

EcoRI will cut Rocky Mountain and North-Central haplotypes; it will not cut Coastal-Cascade or Central Oregon Highland haplotypes

Step 3: Alternative Secondary Digests

For all samples from Step 2 that are either Rocky Mountain or North-Central haplotypes (e.g. cut by EcoRI) proceed using the XmnI and AgeI reaction. For all samples from Step 2 that are either Coastal-Cascade or Central Oregon Highland haplotypes, proceed using the DdeI reaction.

XmnI and AgeI: Differentiates Rocky-Mountain versus North-Central and further differentiations between two North-Central subgroups (See Figure 2.4)

XmnI Recognition Site
GAANN'NNTTC
CTTNN'NNAAG

AgeI Recognition Site
A'CCGGT
TGGCC'A

6 µl reactions consisting of:

0.4 µl Buffer 1 (NEB)
0.033 µl XmnI (NEB:R0101S)
0.13 µl AgeI (NEB: R0552L)
2 µl PCR product (from step 1)
3.38 µl ddH2O
0.06 µl BSA

DdeI: Differentiates Central Oregon Highlands vs Coastal Cascade

DdeI Recognition Site
C'TNAG
GANT'C

6 µl reactions consisting of:

0.4 µl Buffer 3 (NEB)
0.067µl DdeI (NEB:R0175S)
2 µl PCR product from step 1)
3.53 µl ddH2O

Incubation:

37°C for 2 hours

Interpretation of Results:

To tell the difference between Rocky Mountain, North-Central (northern) and North-Central (central) haplotypes:

XmnI will cut Rocky Mountain and North-Central (northern haplotypes) but not North-Central (central haplotyes). AgeI will cut all North-Central haplotypes. Thus, a Rocky Mountain haplotype will have 2 bands (124 and 473 bp), a North-Central (northern) haplotype will have 3 bands (124, 390, 83 bp) and a North-Central (central) haplotype will have 2 bands (514, 83 bp).

To tell the difference between Central Oregon Highlands and Coastal-Cascade:
DdeI will cut the Central Oregon Highlands haplotypes but not Coast-Cascade haplotypes. Thus Coastal-Cascade haplotype will have one band (597 bp) and Central Oregon Highlands haplotypes will have two bands (465 and 132 bp)

A.3 Relative support for different K values in STRUCTURE

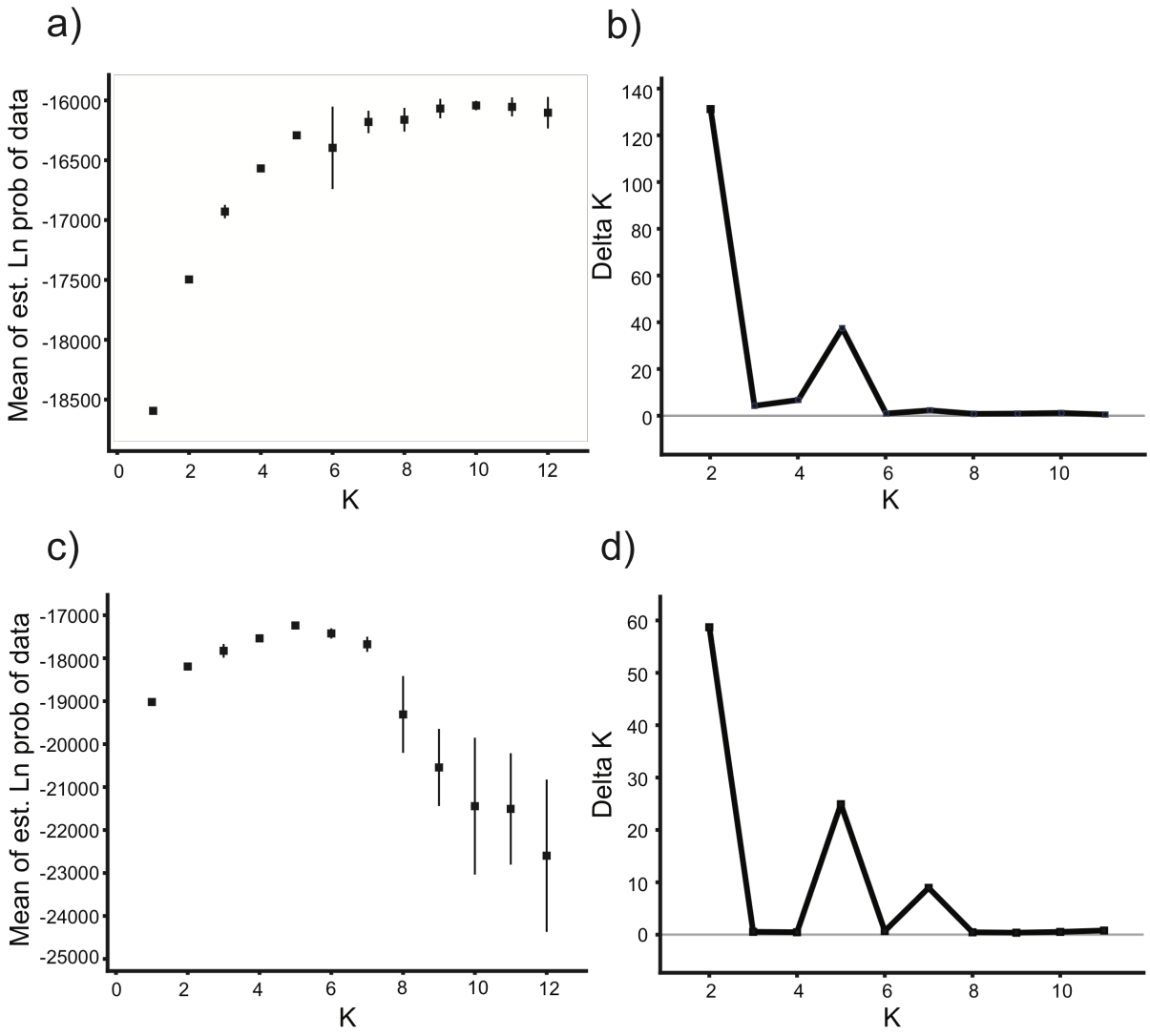

Figure A1. Support for different values of K in STRUCTURE analyses of AFLP data for the long-toed salamander. (a,c) Mean log probability of the data begins to level off at K=5 for 177 and 751 loci respectively. Error bars represent standard deviation. (b,d) Delta K calculated using the method presented in Evanno *et al.* 2005 also shows a peak at K=5 for the analysis of 177 and 751 loci respectively. Plots were produced using the program Structure Harvester (Earl and vonHoldt 2011).

Appendix B Supplementary material for Chapter 3

B.1 Delineating lineage boundaries

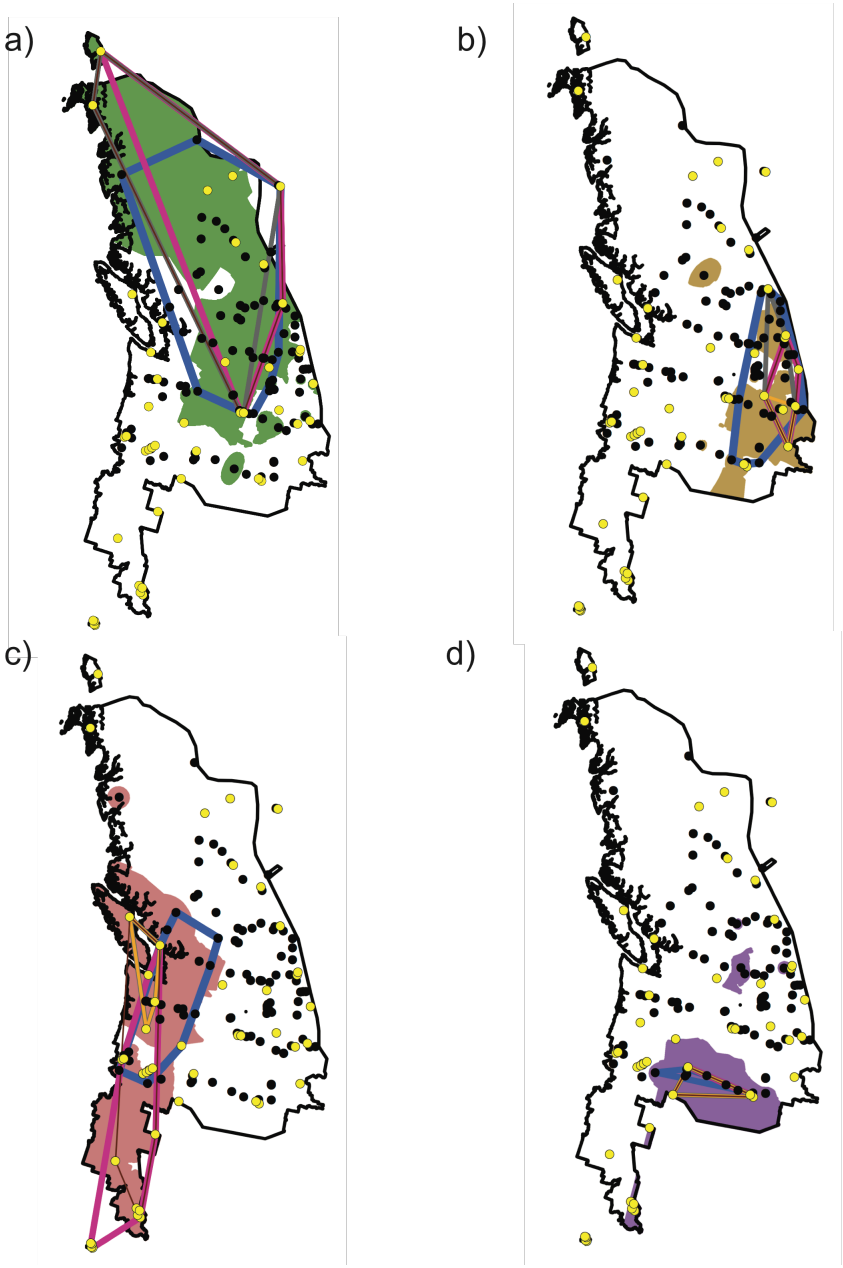

Figure B1. Delineating the boundaries of long-toed salamander lineages: (a) North-Central b) Eastern c) Western d) Oregon-Highlands). Shaded regions show results from IDW interpolation of summarized STRUCTURE ancestry scores from 108 populations (black circles) based on 177 AFLP loci. Yellow points represent sites belonging to two additional, independent datasets: the nuclear sequence dataset of Savage (2008) and the mitochondrial dataset of Lee-Yaw and Irwin (2012). Minimum convex polygons depict clade boundaries based on these genes and are coloured as follows: blue = *cyt b* mtDNA; pink = *gnat2*; grey = *g2i16*; yellow = *col1a1*; brown = *hoxd11*.

B.2 Decisions made when smoothing lineage boundaries

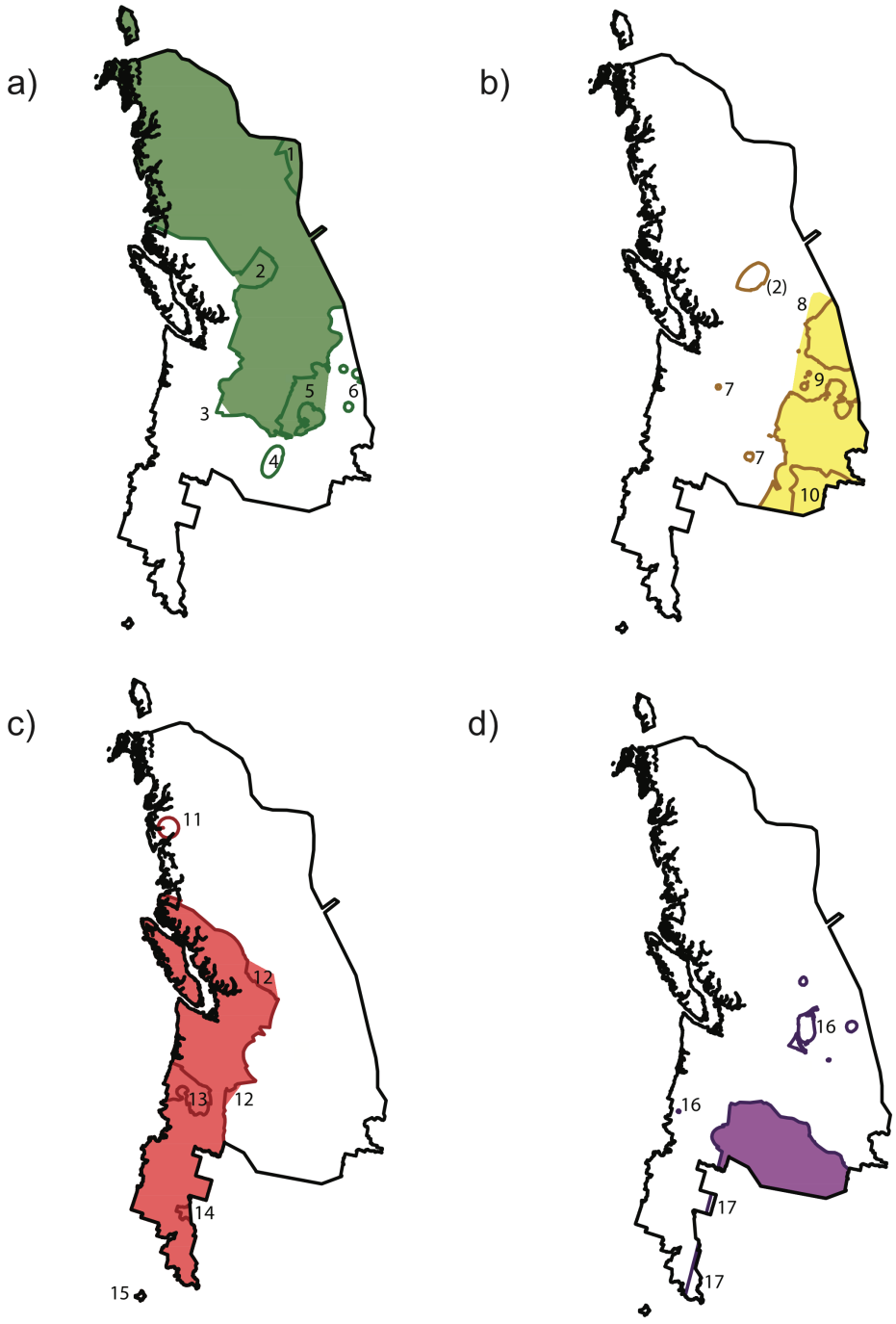

Figure B2. Modifications to long-toed salamander lineage boundaries (solid lines) derived from IDW spatial interpolation of STRUCTURE ancestry scores. Shading indicates final range polygons for the a) North-Central b) Eastern c) Western and d) Oregon-Highlands lineages described by Lee-Yaw and Irwin (2012). Numbers refer to specific changes/decisions made as follows: (North-Central lineage) 1. Extended central range to edge of described species' range in northeast (no other contending lineages in this area). 2.

Reassigned this part of the species' range to the central lineage; original boundaries reflect the ancestry score of a single individual assigned to the eastern lineage; However, the great distance between this individual and the range of the eastern lineage, as well as mtDNA sequence data that places this individual in the North-Central clade, suggests this result likely reflects noise in the AFLP dataset. 3. Range adjusted in this area where AFLP data was sparse to exclude a population with mtDNA suggestive of western ancestry. 4. This polygon removed based on distance away from the central range and because all other data suggests this population is part of the Oregon-Highlands lineage. 5. Boundaries extended to include these disjunct polygons as other populations in the vicinity show evidence of secondary contact; extending the boundaries of this lineage makes the allopatric boundaries of the eastern lineage more conservative. 6.These polygons were removed as they are some distance away from the bulk of the North-Central range and each resulted from a single individual (with <70% of its genome assigned to this lineage); mtDNA in all cases suggests these populations are part of the eastern lineage. (Eastern Lineage) 7. These small polygons were artifacts of the interpolation or reflect single individuals with relatively low eastern ancestry scores (<60%) and were removed based on distance from the rest of the range. 8. Range extended based on mtDNA clade boundaries. 9. Small polygons in this region include populations with mtDNA haplotypes from both the Eastern and North-Central lineages; Eastern range also extended to be continuous. 10. Removed gap in otherwise continuous range; mtDNA from individuals in these regions verifies presence of the eastern lineage. (Western Lineage) 11. Removed this polygon as it reflects an western ancestry score of ~50% from a single individual, is much further north than the rest of the lineage's range and is not supported by mtDNA. 12. Extended the range based on mtDNA clade boundaries. 13. Removed gap in otherwise continuous range; mtDNA from individuals in these gaps verifies presences of western lineage. 14. Extended range to boundaries of species; original omission in this region was an artifact of the boundaries used during interpolation. 15. Removed polygon corresponding to A. m. croceum, an endangered subspecies not included in this study. (Oregon-Highlands lineage). 16. Removed polygons far away from bulk of lineage's range; no other data supports presence of this lineage in these disjunct areas. 17. Eliminated southern tips from lineage boundaries; original inclusion reflects artifact of the boundaries used during interpolation.

B.3 Results from principle components analyses and tests of niche divergence

Table B1. PCA results and tests of niche overlap between long-toed salamander lineages based on climate data.

Lineages Included	Axis	% Variation Explained	Conclusion from Niche Overlap Test	Variable Loadings*								
				tave_sm	tave_sp	tave_at	tave_wt	tmin_sm	tmin_sp	tmin_at	tmin_wt	tmax_sm
All	PC1	65.3	-	0.232	0.241	0.240	0.215	0.219	0.228	0.216	0.205	0.205
	PC2	22.8	-	-0.084	0.016	0.091	0.195	0.018	0.086	0.157	0.213	-0.152
	PC3	5.4	-	0.184	0.163	0.012	-0.102	0.309	0.232	0.156	-0.012	0.059
NC and RM	PC1	70.5	divergence	-0.232	-0.231	-0.229	-0.204	-0.202	-0.217	-0.198	-0.188	-0.221
	PC2	17.9	ns	0.009	0.080	0.112	0.224	0.046	0.134	0.184	0.269	-0.015
	PC3	5.2	conservatism	-0.095	-0.081	-0.019	0.167	-0.428	-0.238	-0.299	0.058	0.116
NC and CC	PC1	64	conservatism	-0.234	-0.246	-0.244	-0.223	-0.229	-0.239	-0.226	-0.213	-0.216
	PC2	28.5	ns	-0.091	0.018	0.073	0.159	0.002	0.086	0.142	0.178	-0.143
	PC3	3.1	divergence	-0.247	-0.155	-0.012	0.148	-0.295	-0.128	-0.033	0.143	-0.194
NC and COH	PC1	75.2	conservatism	-0.225	-0.224	-0.224	-0.203	-0.201	-0.213	-0.198	-0.190	-0.221
	PC2	15.9	conservatism	-0.011	0.071	0.109	0.220	0.028	0.137	0.196	0.266	-0.032
	PC3	4.4	divergence	-0.165	-0.132	-0.015	0.164	-0.401	-0.211	-0.206	0.095	-0.028
RM and COH	PC1	79	conservatism	-0.217	-0.219	-0.221	-0.206	-0.197	-0.210	-0.195	-0.191	-0.211
	PC2	10.2	conservatism	0.016	0.081	0.094	0.212	0.089	0.157	0.231	0.303	-0.036
	PC3	5.9	ns	0.200	0.095	0.043	-0.157	0.304	0.154	0.158	-0.085	0.107
COH and sig	PC1	66.1	conservatism	0.235	0.241	0.240	0.219	0.222	0.234	0.217	0.211	0.221
	PC2	23.5	divergence	0.075	0.012	-0.083	-0.168	0.011	-0.038	-0.144	-0.171	0.121
	PC3	4.5	divergence	0.166	0.127	0.007	-0.108	0.331	0.225	0.172	0.015	0.011
COH and CC	PC1	62.5	conservatism	-0.243	-0.244	-0.239	-0.197	-0.224	-0.216	-0.181	-0.170	-0.221
	PC2	30.9	ns	-0.057	0.076	0.113	0.217	0.089	0.173	0.233	0.254	-0.147
	PC3	2.9	divergence	0.264	0.099	0.038	-0.175	0.348	0.112	0.079	-0.100	0.172
CC and sig	PC1	64.5	conservatism	-0.242	-0.241	-0.243	-0.226	-0.218	-0.217	-0.205	-0.202	-0.223
	PC2	26.5	ns	0.032	-0.065	-0.073	-0.148	-0.105	-0.153	-0.192	-0.207	0.122
	PC3	3.5	divergence	-0.105	0.067	-0.075	0.005	0.144	0.220	0.149	0.145	-0.262

Lineages Included	Axis	Variable Loadings*								
		tmax_sp	tmax_at	tmax_wt	mcmt	mvmt	mat	Td	nffd	shm
All	PC1	0.234	0.234	0.213	0.201	0.231	0.243	-0.038	0.215	0.212
	PC2	-0.051	0.017	0.160	0.220	-0.081	0.069	-0.356	0.145	-0.150
	PC3	0.090	-0.118	-0.198	-0.160	0.159	0.058	0.324	0.194	-0.267
NC and RM	PC1	-0.228	-0.224	-0.207	-0.192	-0.232	-0.231	-0.061	-0.202	-0.217
	PC2	0.033	0.044	0.157	0.236	0.008	0.110	-0.337	0.128	-0.150
	PC3	0.041	0.178	0.278	0.247	-0.068	-0.011	-0.405	-0.349	0.113
NC and CC	PC1	-0.239	-0.245	-0.228	-0.216	-0.234	-0.247	0.086	-0.222	-0.183
	PC2	-0.044	0.010	0.132	0.172	-0.091	0.056	-0.314	0.131	-0.237
	PC3	-0.172	0.006	0.153	0.187	-0.228	-0.049	-0.409	-0.058	0.223
NC and COH	PC1	-0.223	-0.222	-0.208	-0.195	-0.225	-0.225	-0.030	-0.201	-0.217
	PC2	0.019	0.047	0.162	0.238	-0.008	0.102	-0.384	0.135	-0.134
	PC3	-0.070	0.098	0.232	0.215	-0.139	-0.034	-0.539	-0.266	0.109
RM and COH	PC1	-0.215	-0.217	-0.208	-0.196	-0.218	-0.221	-0.058	-0.202	-0.216
	PC2	0.019	-0.014	0.100	0.237	0.014	0.102	-0.323	0.166	-0.088
	PC3	0.047	-0.042	-0.217	-0.240	0.187	0.048	0.654	0.171	-0.096
COH and sig	PC1	0.233	0.234	0.214	0.206	0.232	0.244	0.019	0.222	0.210
	PC2	0.056	-0.017	-0.152	-0.205	0.077	-0.045	0.347	-0.126	0.163
	PC3	0.033	-0.144	-0.238	-0.132	0.190	0.051	0.352	0.216	-0.257
COH and CC	PC1	-0.244	-0.245	-0.216	-0.184	-0.238	-0.243	-0.040	-0.184	-0.195
	PC2	-0.017	-0.014	0.152	0.235	-0.072	0.096	-0.319	0.237	-0.217
	PC3	0.080	-0.005	-0.246	-0.202	0.296	0.054	0.503	-0.008	-0.071
CC and sig	PC1	-0.241	-0.237	-0.231	-0.221	-0.240	-0.245	0.008	-0.203	-0.185
	PC2	0.015	0.030	-0.075	-0.158	0.041	-0.069	0.312	-0.195	0.244
	PC3	-0.072	-0.239	-0.140	-0.024	-0.142	-0.023	-0.209	0.234	-0.167

Lineages Included	Axis	Variable Loadings*						
		pas	ppt_sm	ppt_sp	ppt_at	ppt_wt	map	msp
All	PC1	-0.219	-0.186	-0.116	-0.080	-0.065	-0.105	-0.187
	PC2	0.084	0.086	0.346	0.382	0.378	0.369	0.164
	PC3	-0.022	0.494	0.005	0.060	-0.069	0.045	0.409
NC and RM	PC1	0.205	0.197	0.140	0.157	0.131	0.166	0.194
	PC2	0.190	0.151	0.340	0.327	0.352	0.329	0.208
	PC3	0.013	-0.253	0.102	-0.153	-0.014	-0.078	-0.180
NC and CC	PC1	0.208	0.170	0.028	0.063	0.018	0.059	0.145
	PC2	0.119	0.215	0.362	0.355	0.359	0.360	0.270
	PC3	-0.063	-0.448	-0.068	-0.020	0.090	-0.050	-0.384
NC and COH	PC1	0.209	0.207	0.137	0.174	0.144	0.176	0.203
	PC2	0.147	0.119	0.381	0.302	0.355	0.309	0.178
	PC3	-0.089	-0.225	-0.015	-0.173	-0.114	-0.147	-0.197
RM and COH	PC1	0.208	0.200	0.191	0.175	0.152	0.187	0.204
	PC2	0.172	0.084	0.298	0.365	0.403	0.326	0.125
	PC3	0.103	0.242	0.071	0.111	0.094	0.128	0.202
COH and sig	PC1	-0.227	-0.198	-0.057	-0.045	-0.023	-0.049	-0.184
	PC2	-0.084	-0.046	-0.389	-0.394	-0.388	-0.392	-0.183
	PC3	-0.009	0.451	0.060	0.045	-0.044	0.036	0.414
COH and CC	PC1	0.235	0.172	0.137	0.141	0.131	0.142	0.165
	PC2	0.028	0.230	0.291	0.290	0.291	0.290	0.252
	PC3	0.304	0.194	0.144	0.144	0.156	0.154	0.182
CC and sig	PC1	0.234	0.163	0.080	0.107	0.076	0.097	0.151
	PC2	-0.039	-0.245	-0.342	-0.338	-0.335	-0.341	-0.289
	PC3	-0.217	0.320	-0.343	-0.192	-0.396	-0.281	0.167

B.4 Suitability maps when contact zones are included in niche models

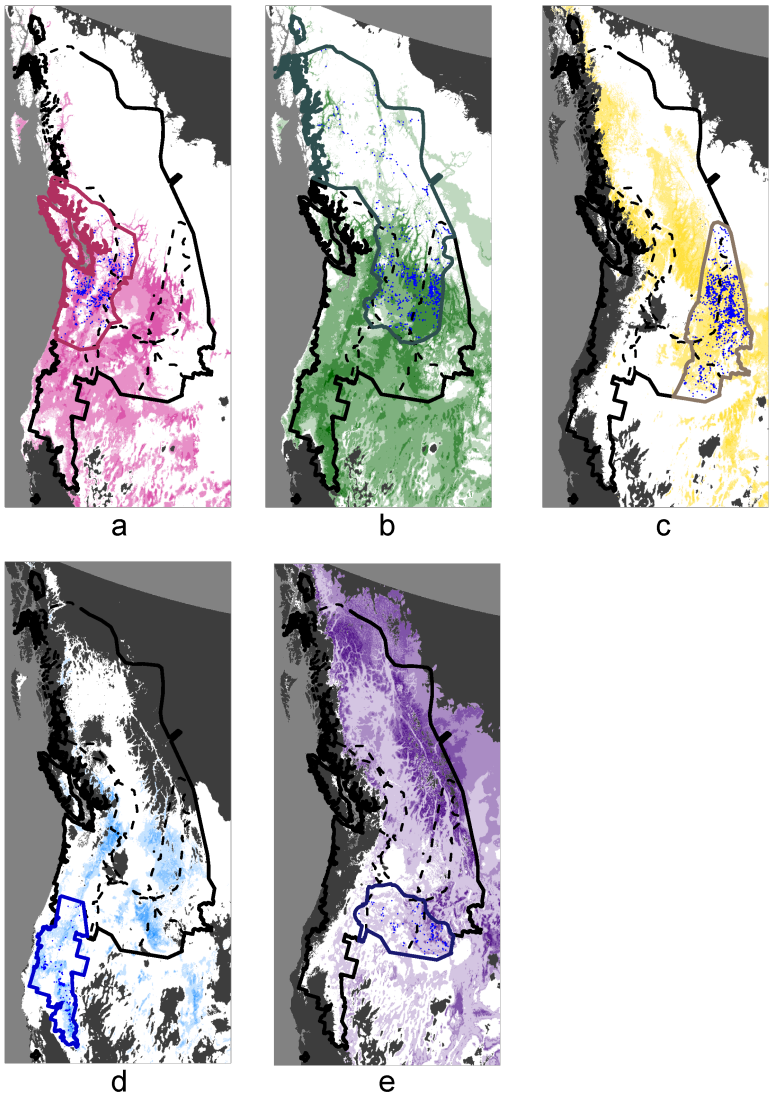

Figure B3. Suitability maps when contact zones are included in niche model calibration for a) Coastal-Cascade b) North-Central c) Rocky Mountains d) *A. m. sigillatum* e) Central-Oregon Highlands long-toed salamanders. Darker colours indicate higher suitability. All areas where suitability was lower than 95% of the focal lineage's range were assigned a value of zero (white) for plotting purposes. Small blue dots represent sampling locations. Dark grey areas correspond to places where the range of one or more climatic variables was outside that used to calibrate the model and thus where model extrapolation would have been necessary to make predictions (*e.g.* results from the MESS analysis). The solid black line delimits the boundaries of *A.macrodactylum* with dashed lines showing the lineage boundaries as per Figure 3.2. Solid coloured lines delineate the allopatric range of the focal lineage.

B.5 Suitability across lineage boundaries when niche models include contact zones

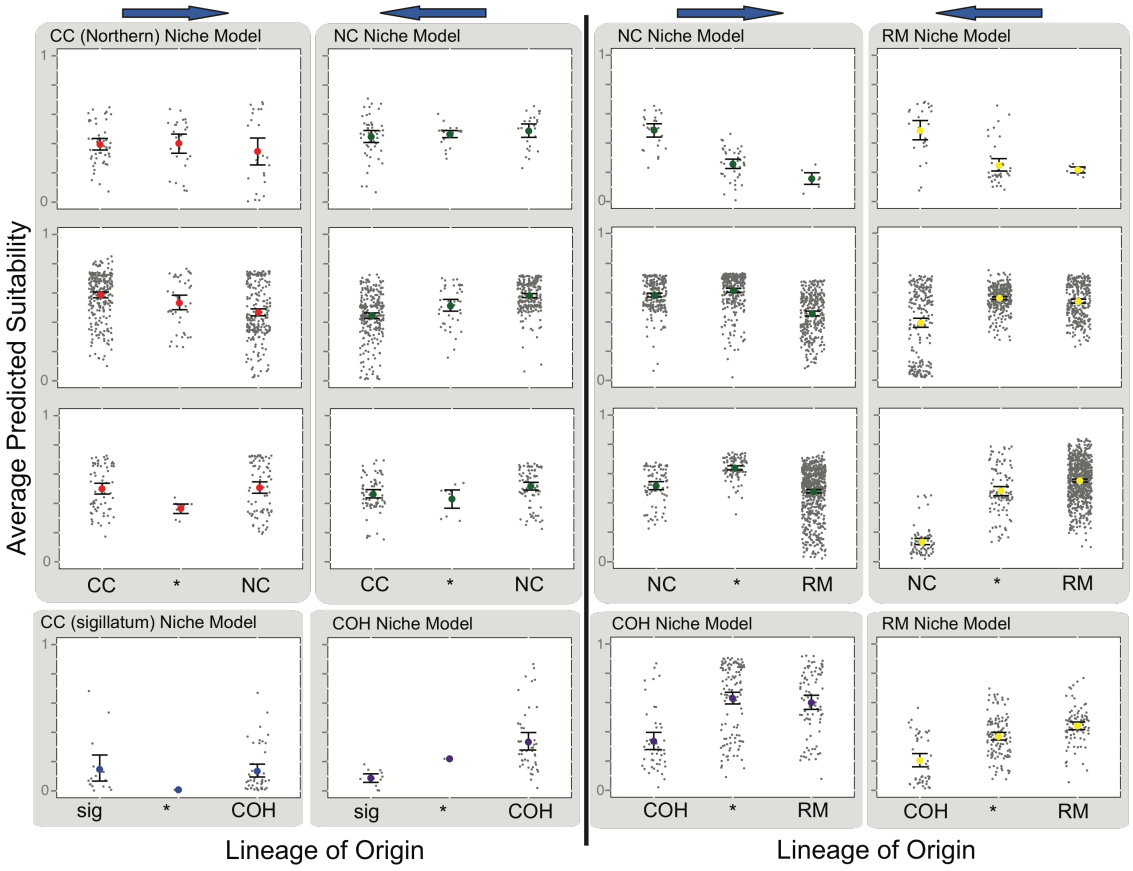

Figure B4. Average climatic suitability (with 95% CI based on 1000 bootstrap replicates) of populations (shown as grey dots in plots) across range boundaries when contact zones are included in model calibration. The panels to the left of the line show results for the parapatric boundary between the Coastal-Cascade and North-Central (NC) and *A. m. sigillatum* (sig) and Central-Oregon Highlands (COH) lineages. The panels to the right of the line show results for the parapatric boundary between the North-Central and Rocky Mountains (RM) and Central-Oregon Highlands and Rocky Mountains lineages. Contact zones in the comparisons are denoted with a star. In both sets of panels, the graphs to the left are based on the niche model of the western-most lineage in the comparison and the graphs to the right are based on the niche model of the eastern-most lineage in the comparison (also indicated by the colour of the average values in each plot). Plots are grouped to reflect the different ANOVA analyses (Table 3).

Appendix C Supplementary material for Chapter 4

C.1 PCR-RFLP protocols for nuclear genes that diagnose NC and RM

For all PCR reactions the following 25 µl reaction mix was used:

0.25 µl dNTPs (10 mM)
2.5 µl 10x buffer (New England BioLabs)
0.75 µl MgCl (50 mM)
1 µl R primer (10 µM)
1 µl F primer (10 µM)
0.2µl TAQ
18.4 µl ddH2O
1.0 µl DNA

For all PCR reactions the following program was used:
94°C for 2 mins
35 cycles of:
{94°C for 30 s
TA for 45 s
72°C for 1 min 10 s}
72°C for 7 mins

All PCR products were digested at 37°C for 2 hours

col1a1

R primer: 5' –TCG TTT TTG GAG GTG TAA TGC – 3'
F primer: 5' –AAA TCT CTG CAT TCT CCC AGA – 3'
TA = 60°C

NC and RM individuals can be distinguished on agarose gel from the PCR products due to a 522 bp deletion in NC individuals.

hoxd11

R primer: 5' –AAC CAG GTC CCC TTA GTT CTA TTC AGG – 3'
F primer: 5' –ACA TCA TCT TCG GGA CTG TAA CAA GG – 3'
TA = 55°C

Digest reaction (6 µl):
0.4 µl EcoRI buffer (New England BioLabs)
0.134 µl PsiI (New England BioLabs)
2 µl PCR product from step 1
3.47 µl ddH2O

NC and RM individuals can be distinguished on agarose gel from digest product as PsiI cuts RM but not NC individuals.

g2-i16

R primer: 5' –TGA GGT GCA TGA TGG TCT TTG – 3'
F primer: 5' – CTT CCT TCC TTC GAG CTG GTG TCT A – 3'
TA = 62°C

Digest reaction (6 µl):
0.4 µl EcoRI buffer (New England BioLabs)
0.033 µl MspI (New England BioLabs)
2 µl PCR product from step 1
3.57 µl ddH2O

NC and RM individuals can be distinguished on agarose gel from digest product as MspI cuts NC but not RM individuals.

C.2 Information about AFLP primers used in contact zone

Table C1. AFLP primer combinations used to survey individuals in a contact zone between NC and RM long-toed salamanders.

EcoRI primer (*NNN-3')	MseI primer (†NNN-3')	Dye	Number of Polymorphic Fragments (after filtering)
AGCA	CATG	NED	49
AGCG	CATA	PET	51
AGCT	CATA	VIC	43
AGCC	CATC	FAM	39

EcoRI primer: GACTGCGTACCAATTC*
MseI primer: GATGAGTCCTGAGTAA†