Title
The Dividing Link: Speciation and Hybridization in the Salamander Ring Species *Ensatina eschscholtzii*

Author
Devitt, Thomas James

Publication Date
2010

Peer reviewed|Thesis/dissertation

The Dividing Link: Speciation and Hybridization in the Salamander Ring Species *Ensatina eschscholtzii*

By

Thomas James Devitt

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Craig Moritz, Co-chair
Professor Jimmy McGuire, Co-chair
Professor David B. Wake
Professor George K. Roderick

Spring 2010

Abstract

The Dividing Link: Speciation and Hybridization in the Salamander Ring Species *Ensatina eschscholtzii*

by

Thomas James Devitt

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Craig Moritz, Co-chair

Professor Jimmy McGuire, Co-chair

Plethodontid salamanders of the *Ensatina eschscholtzii* complex have received special attention from evolutionary biologists because they represent one of the very few examples of a ring species, a case where two reproductively isolated forms are connected by a chain of intergrading populations surrounding a central geographic barrier. *Ensatina* has become a textbook example of speciation, yet there still remain fundamental gaps in our knowledge of this fascinating system. In this study, consisting of three components, I extend previous work on the *Ensatina* complex in new directions.

In Chapter 1, I conducted a fine-scale genetic analysis of a hybrid zone between the geographically terminal forms of the ring using Bayesian methods for hybrid identification and classification in combination with mathematical cline analyses. F1s and pure parentals dominated the sample. Cline widths were concordant and narrow with respect to dispersal, but there is cytonuclear discordance, both in terms of introgression and the geographic position of mitochondrial versus nuclear clines. Nearly all hybrids possess mitochondrial DNA from one parental type (*klauberi*) suggesting isolation is asymmetrical. Selection against hybrids is inferred to be strong (~21%), but whether this selection is endogenous (genetically-based) or exogenous (environmentally-based) remains to be tested.

In Chapter 2, I investigated the role of late Quaternary climate change on phylogeographic patterns within the Large-blotched Ensatina (*Ensatina eschscholtzii klauberi*). Intersecting species distribution models constructed under current climatic conditions, as well as two different historical time periods (Last Glacial Maximum, 21 ka and mid-Holocene, 6 ka), predicted stable refugial areas where the species may have persisted throughout climatic fluctuations. Significant phylogeographic structure exists, but geographic structuring of genetic variation by refugia was not supported. Results suggest that populations in putative refugia have

not been isolated for very long, or that gene flow may have masked any earlier periods of divergence in allopatry.

In Chapter 3, I conducted a multilocus phylogeographic analysis of the entire *Ensatina eschscholtzii* complex to reexamine previous phylogenetic hypotheses based on mitochondrial DNA alone. A concatenation approach was used in addition to newer methods that model the relationship between the species tree and the gene trees embedded within them. The concatenated tree was similar to previous mitochondrial trees, identifying well-supported coastal and inland clades, and recovering *oregonensis* and *platensis* as paraphyletic. The concatenated tree was not well resolved at the base. Basal relationships recovered by the species tree were well resolved, but most relationships were not well supported compared to the concatenated tree. Results are generally consistent with previous efforts based on mtDNA, but provide further resolution to the *Ensatina* phylogeny, while highlighting the difficult nature of inferring species trees from samples of closely related populations that are experiencing gene flow.

Dedication

I dedicate this work to my parents, Diane L. Devitt and Robert G. Devitt (1930-2007). Thank you for your love, support, and for encouraging me to follow my dreams. I couldn't have done it without you. I wish you could have been here to see me graduate, Dad. I hope I've made you proud.

Table of Contents

Introduction

Evolutionary biologists have long been fascinated with the origin, development, and maintenance of geographic variation in nature. Initially, complete geographic isolation was thought to be necessary for speciation to occur (Huxley 1942). Only in recent decades has it been demonstrated that intense geographic differentiation and speciation can occur among continuously distributed populations isolated by distance alone, particularly when subject to divergent ecological selection (Endler 1977; Coyne and Orr 2004; Schluter 2009). The role of gene flow is critical in population differentiation because gene flow counteracts divergence among populations. Thus, factors that reduce gene flow increase genetic isolation thereby promoting population differentiation and speciation. A number of factors influence the relative role of gene flow in speciation, depending upon whether populations have diverged incidentally through genetic drift (Gavrilets 2003), or because of selective forces such as local adaptation (Ehrlich and Raven 1969; Endler 1973; Slatkin 1987).

Much of speciation research has focused on modes of speciation involving restricted gene flow (i.e., allopatric and parapatric speciation). Stejneger (in Jordan 1905) was the first to speculate on one of the most interesting geographic scenarios of speciation in which restricted gene flow was thought to play a pivotal role. He suggested that cases may exist in nature where a single species expands around a central geographic barrier along two separate pathways, with the two terminal forms gradually diverging and eventually becoming reproductively isolated where they meet and reconnect on the other side of the barrier. Cain (1954) coined the term "ring species" to describe this scenario, and Mayr (1942) called such instances of "circular overlap" the "perfect demonstration of speciation," citing a number of potential cases in nature. Although Mayr (1942) originally discussed both historical divergence and the connectivity of neighboring populations united by gene flow as causes of circular overlaps, he later deemphasized the role of gene flow, due to a lack of evidence (Mayr 1970). In contrast, Dobzhansky placed particular emphasis on the role of gene flow as a "genetic bridge" uniting a chain of interbreeding populations in his work on circular overlaps in *Drosophila paulistorum* (Dobzhansky and Spassky 1959; Dobzhansky et al. 1964; Dobzhansky and Pavlovsky 1967; Mayr 1970) and his discussion of the *Ensatina* complex (Dobzhansky 1958). While some stress smooth gene flow around the ring as a requirement for ring species status (Highton 1998; Coyne and Orr 2004), others emphasize historical range shifts with multiple separation and reconnection events, placing little significance on contemporary gene flow as a genetic bridge uniting neighboring populations (Wake and Schneider 1998; Pereira and Wake 2009). Depending on which feature emphasis is placed, disagreement may arise surrounding which cases are actually ring species (Irwin et al. 2001).

Two features make ring species ideal systems for speciation research (Irwin et al. 2001). First, ring species allow the history and mechanisms of species formation to be traced back in time through the geographically differentiated populations connecting the reproductively isolated forms. Second, ring species show in a novel way that reduced gene flow promotes speciation (Irwin et al. 2001). Irwin et al. (2001) suggested that the role of gene flow in slowing divergence of the terminal forms of a ring species depends upon whether populations are subject to divergent ecological selection. They predict that ring species may form in the absence of divergent ecological selection only when gene flow between terminal forms is low. Conversely, if ecological conditions vary around the ring and populations are subject to diversifying ecological selection, ring species may readily form, even with considerable gene flow (Irwin et

al. 2001). Although no case is perfect, Irwin et al. (2001) found that two cases best exhibited these features, the Greenish warbler complex of Asia (*Phylloscopus trochiloides*) (Ticehurst 1938; Irwin 2000; Irwin et al. 2001; Irwin and Irwin 2002) and the *Ensatina eschscholtzii* complex (Stebbins 1949; Dobzhansky 1958; Moritz et al. 1992; Wake 1997).

In his detailed analysis of geographic variation and speciation in *Ensatina* Stebbins (1949) concluded that the four previously described species of *Ensatina* (*eschscholtzii*, *sierrae*, *croceater*, and *platensis*) actually represented a single polytypic species made up of seven well-defined subspecies, based primarily on differences in color pattern. He hypothesized that the complex originated in northwestern North America, perhaps, although not necessarily, in the Klamath Range region of northwestern California and southwestern Oregon. There, an ancestral population divided into two subpopulations that expanded southward around the arid Central Valley of California along two separate paths, one along the relatively low-elevation coast ranges, the other inland along the western slopes of the higher-elevation Sierra Nevada and its foothills. Eventually, the two lineages came back into contact in southern California, creating a ring encircling the Central Valley. Three of the coastal subspecies (*oregonensis*, *xanthoptica* and *eschscholtzii*) evolved uniform, unblotched color patterns. The fourth coastal form, *picta*, inhabiting a narrow coastal strip in extreme northwestern California and southwestern Oregon, possessed a pattern characterized by irregular dark blotches and had already been described as a separate subspecies based on its unique color pattern (Wood 1940). Stebbins thought *picta* closely resembled the ancestral type because it possessed the most generalized color pattern, characteristics of which could be seen in the other six subspecies. The inland, montane lineages (*platensis*, *croceater* and *klauberi*) evolved bold, distinctly blotched color patterns. Stebbins hypothesized that the striking color pattern variation exhibited by the seven subspecies evolved as different strategies for predator avoidance: crypsis via background matching in *oregonensis*, *picta*, and *eschscholtzii*, crypsis via disruptive coloration in *platensis*, *croceater* and *klauberi*, and mimicry in *xanthoptica*, which bears an unmistakable resemblance to the highly toxic newts of the genus *Taricha*, an observation supported by experimental evidence (Kuchta et al. 2008).

Stebbins (1949) described broad (up to 150 km) intergrade zones where neighboring subspecies met, except where the two ends of the ring come together in the Peninsular Range region of southern California. Although the two geographically terminal forms, *klauberi* and *eschscholtzii*, had not been found in sympatry when Stebbins first published his biogeographical hypothesis in 1949, he had found them in close proximity and expected that they would be found to co-occur. Indeed, the terminal forms were later found in sympatry in four geographically isolated contact zones; limited hybridization was discovered at three of them, while complete reproductive isolation occurred at the fourth, southernmost contact zone (Stebbins 1957; Brown and Stebbins 1964; Brown 1974; Wake et al. 1986).

Decades of genetic work by D. B. Wake and colleagues have uncovered remarkably high, geographically structured genetic diversity within *Ensatina*, suggesting the complex is old and has featured periods of geographic isolation and multiple instances of secondary contact. Allozymes have revealed high genetic distances both within and between subspecies (Wake and Yanev 1986; Jackman and Wake 1994; Pereira and Wake 2009). Mitochondrial DNA gene genealogies revealed that two of the subspecies (*oregonensis* and *platensis*) are paraphyletic and contain multiple lineages (Moritz et al. 1992; Kuchta et al. 2009a; Kuchta et al. 2009b). Overall however, these studies are generally consistent with Stebbins' (1949) biogeographic hypothesis in that genetic distances are higher between the geographically terminal populations than most adjacent populations (Moritz et al. 1992; Pereira and Wake 2009), and within subspecies there is

a positive correlation between genetic and geographic distance showing that distance is in fact a barrier to gene flow (Jackman and Wake 1994).

In this study, consisting of three primary components, I extend previous work on the *Ensatina* complex in new directions. In Chapter 1, I conduct a fine scale genetic analysis of a hybrid zone between the geographically terminal forms of the ring, *eschscholtzii* and *klauberi*, to provide insight into evolutionary processes at this critical juncture. In Chapter 2, I examine the role of recent (late Quaternary) climate change in driving patterns of population connectivity and divergence in the Large-blotched Ensatina (*Ensatina eschscholtzii klauberi*) using a combination of paleomodeling, landscape genetic, and phylogeographic approaches. Finally, in Chapter 3, I reexamine phylogenetic relationships within the *Ensatina* complex with multilocus DNA sequence data using new approaches that explicitly model the relationship between species trees and the gene trees embedded within them.

Acknowledgements

I owe it to a lot of people for helping me get through my Ph.D. It's been a long road, and I couldn't have done it without them.

For support and guidance throughout the course of this project, I am indebted to my co-advisors, Jim McGuire and Craig Moritz, and my unofficial third advisor, Dave Wake.

I've known Jim since I was an undergraduate at the University of Texas. Along with Adam Leaché, I was Jim's first graduate student at LSU, where I did a master's with him before coming to Berkeley. Jim has been an amazing advisor, mentor, and friend to me. He gives the perfect amount of guidance without micromanaging. He's *always* willing to go to bat for his students, no matter what. Jim believed in me even when I didn't believe in myself. Having that kind of support makes all the difference. And I'll never forget our field trips to Indonesia and Baja California. Jim was the only one strong enough to lift that giant slab of granite in the Sierra San Pedro Mártir where he found the only *klauberi* of the entire trip — only the 3rd specimen ever collected in Baja California! I'm proud to say I was one of Jim's students.

I really had the best of both worlds having Craig as a co-advisor. Craig has taught me a lot, especially to never give up. I still can't believe he was willing to write and submit an NSF grant with me *four times* before we finally got funded. Craig didn't need that grant; he did it to help me free up my time from teaching so that I could spend the months in the field necessary for my hybrid zone study. Craig really challenged me, and I'm a better scientist for it. I always leave Craig's office feeling more confident, motivated, and excited about my work. And who else but Craig would hold the biannual lab retreat at my study site just to help me collect samples. Craig and about 12 people from the lab drove all the way down from Berkeley to Palomar Mountain (about a 9-hour drive) just to lend a hand in the field. I'll never forget that. I feel so lucky to have worked with him.

Although not technically a co-advisor on paper (turns out they only let you have two), I consider Dave my third advisor. He's a mentor in every sense of the word. I'll never forget the first time I met him. It was at the herp meetings in 2003. I hadn't yet started at Berkeley, but Dave knew that I was coming as Jim's student in the fall. I'd seen Dave at the meeting but never had the courage to go up and talk to him. At one point during a session break, Dave saw my nametag and actually came up to me and said, "Oh hi, you're Tom Devitt. I'm David Wake. We're really looking forward to having you at the MVZ." I was floored. Dave *Wake* came up and introduced himself to *me*! I'm so glad that I decided to work on *Ensatina* because it enabled me to get to know Dave. His passion and enthusiasm for science, especially involving *Ensatina*, is contagious. My dissertation was greatly improved by having Dave on my committee.

I would also like to thank my "outside" committee member, George K. Roderick, for serving on my committee and reading my dissertation, even though I didn't give him much time. I learned a lot from George in his Population Genetics seminar and he was always willing to give advice when I needed it.

I learned so much from my fellow grad students and postdocs in the MVZ during my tenure at Berkeley. For offering to read my DDIG proposal(s), helping me run analyses (especially you, Fujita), and so much more, I owe them my sincerest thanks. I've made some lifelong friends in the MVZ, and it wouldn't have been the same anywhere else.

I am especially grateful to my family, particularly my sister Carolyn, for making me laugh and keeping me sane. And last, but certainly not least, I thank the best thing that's ever happened to me, my wife Susan, for supporting me all along the way.

Thomas James Devitt

CURRICULUM VITAE

Museum of Vertebrate Zoology
3101 Valley Life Sciences Building
University of California
Berkeley, CA 94720-3160

Telephone: (510) 325-2694
Fax: (510) 643-8238
Email: tdevitt@berkeley.edu

PROFESSIONAL PREPARATION

- University of California, Berkeley, August 2003 – May 2010; Degree: Ph.D. in Integrative Biology; advisors: Craig C. Moritz and Jimmy A. McGuire
- Louisiana State University, Baton Rouge, Louisiana, August 2000 – May 2003; Degree: Master of Science in Biological Sciences; advisor: Jimmy A. McGuire
- The University of Texas, Austin; January 1996 - December 1999; Degree: Bachelor of Science in Biology; Major: Ecology, Evolution, and Conservation
- The University of California at Davis; September 1995 – December 1995 (transferred to the University of Texas)

RESEARCH INTERESTS

- Speciation, hybridization, reinforcement, sexual selection, phylogeography

PUBLICATIONS

Devitt, T. J., Pereira, R., Jakkula, L., Alexandrino, J., Bardeleben, C. and C. Moritz. 2009. Isolation and characterization of 15 polymorphic microsatellites in the Plethodontid salamander *Ensatina eschscholtzii*. Molecular Ecology Resources 9:966-969.

Devitt, T. J. and T. J. LaDuc. 2008. The *Trimorphodon biscutatus* (Squamata: Colubridae) species complex revisited: a multivariate statistical analysis of geographic variation. Copeia 2008 (2):370-387.

Devitt, T. J. and S. E. Cameron. 2007. "Geographic Distribution: *Masticophis taeniatus*." Herpetological Review 38(2): 221.

Devitt, T. J., B. I. Crother, A. Meier, F. Burbrink, and J. Boundy. 2007. "Maximum length: *Pituophis catenifer*." Herpetological Review 38(2): 209-210.

Devitt, T. J. 2006. Phylogeography of the Western Lyresnake (*Trimorphodon biscutatus*): Testing Aridland Biogeographic Hypotheses Across the Nearctic-Neotropical Transition. *Molecular Ecology* 15:4387-4407.

Heim, C. D., B. Alexander, R. W. Hansen, J. H. Valdez-Villavicencio, **T. J. Devitt**, B. D. Hollingsworth, J. A. Soto-Centeno, and C. R. Mahrdt. 2005. "Geographic Distribution: *Ensatina eschscholtzii klauberi*." Herpetological Review 36(3): 330-331.

Devitt, T. J., R. I. Hill, and S. E. Cameron. 2005. "Diet: *Charina bottae*." Herpetological Review 36(2): 189-190.

APPOINTMENTS

February 2007-May 2010
Graduate student researcher in Department of Integrative Biology, University of California, Berkeley.

January 2003 – May 2004; August 2005 – December 2005
Graduate student instructor in Department of Integrative Biology, University of California, Berkeley. Courses taught include Natural History of the Vertebrates and General Biology ("Field Section").

July 2004 – June 2005
Graduate student fellow in NSF-funded GK-12 program "Exploring California Biodiversity". Developed a curriculum for two advanced placement high school environmental science classes using facilities and resources of the Berkeley Natural History Museums.

August 2000 – May 2001; August – December 2002
Teaching assistant in the Department of Biological Sciences, Louisiana State University. Courses taught include a laboratory course entitled Basic Biological Investigations, Herpetology Laboratory, and Evolution laboratory.

September 2001 – May 2002
Curatorial assistant, Museum of Natural Science, Louisiana State University.
Assisted with general maintenance of the herpetology collection.

November 1997 - May 1998; June 1999 – May 2000
Curatorial Assistant, Texas Natural History Collection of the Texas Memorial Museum.
Assisted with general maintenance of the herpetology collection.

RESEARCH GRANTS

NSF Doctoral Dissertation Improvement Grant (DDIG) 2009 ($14,908)
Howie Wier Conservation Award (Anza Borrego Foundation) December 2007 ($2,000)
UC MEXUS, June 2006 ($1,500)
Martens Fund, Museum of Vertebrate Zoology, May 2006 ($2,000)
Mildred E. Mathias Graduate Student Research Grant, January 2006 ($700)
Martens Fund, Museum of Vertebrate Zoology, May 2005 ($1,300)
Sigma Xi Grant in Aid of Research, June 2005 ($500)
AMNH Theodore Roosevelt Memorial Fund Award, May 2004 ($1,700)
Martens Fund, Museum of Vertebrate Zoology, May 2004 ($2,000)
Louise Kellogg Fund, Museum of Vertebrate Zoology, May 2003 ($2,000)
Sigma Xi Grant in Aid of Research, May 2002 ($300)
SSAR Grant-in-Herpetology for Conservation, April 2002 ($500) – declined
Charles Stearns Grant, California Academy of Sciences, June 2001 ($500)

TEACHING AWARDS

Outstanding Graduate Student Instructor 2005-2006, UC Berkeley Department of Integrative Biology

ORAL PRESENTATION AWARDS

HL Award for Graduate Research Honorable Mention, 2008 JMIH
ASIH Stoye Award for General Herpetology, 84th Annual Meeting, ASIH, May 2004
HL Jaeger Award for Graduate Research Honorable Mention, 2002 JMIH

PRESENTATIONS

Devitt, T. J., Cameron, S., and M. Koo. "Hybridization and Ecological Differentiation Between Terminal Forms of the Salamander Ring Species *Ensatina eschscholtzii*." Joint Meeting of Ichthyologists and Herpetologists 2009, Portland, OR

Devitt, T. J., Cameron, S., and M. Koo. "Hybridization and Ecological Differentiation Between Terminal Forms of the Salamander Ring Species *Ensatina eschscholtzii*." Evolution 2009, Moscow, ID

Devitt, T. J. "Asymmetrical Reproductive Isolation Between Terminal Forms of the Ring Species *Ensatina eschscholtzii*." Joint Meeting of Ichthyologists and Herpetologists 2008, Montreal, QC, Canada

Devitt, T. J. "Asymmetrical Reproductive Isolation Between Terminal Forms of the Ring Species *Ensatina eschscholtzii*." Evolution 2008, St. Paul, MN

Devitt, T. J. "An Integrative Approach to Understanding Speciation in the Ring Species *Ensatina eschscholtzii*" 5th Conference on the Biology of Plethodontid Salamanders, August 2007, San Cristóbal de las Casas, Chiapas, Mexico

Devitt, T. J. and R. Hijmans. "The *Ensatina eschscholtzii* Complex Revisited: Exploring the Potential for Ecological Isolation Using Niche Modeling" Joint Meeting of Ichthyologists and Herpetologists 2005, Tampa, FL

Devitt, T. J.. "Phylogeography of Lyresnakes (*Trimorphodon*): Testing Biogeographic Hypotheses for North and Middle American Aridlands" Joint Meeting of Ichthyologists and Herpetologists 2004, Norman, OK

Devitt, T. J. "Testing the Hypothesis of Multiple Lineages Within *Trimorphodon biscutatus* Using Morphometric and Mitochondrial DNA Sequence Data" Evolution 2002, Urbana-Champaign, IL

Devitt, T. J. "Testing the Hypothesis of Multiple Lineages Within Trimorphodon biscutatus Using Morphometric and Mitochondrial DNA Sequence Data" Joint Meeting of Ichthyologists and Herpetologists 2002, Kansas City, MO

PROFESSIONAL MEMBERSHIPS

Society for the Study of Evolution, Society of Systematic Biologists, American Society of Ichthyologists and Herpetologists, Herpetologists' League, Society for the Study of Amphibians and Reptiles

SYNERGISTIC ACTIVITIES

Reviewer for: Molecular Ecology, Herpetologica, Herpetological Monographs, Herpetological Review, Journal of Herpetology

Chapter 1

The Dividing Link: Genetic Analysis of a Hybrid Zone Between Terminal Forms of the Salamander Ring Species *Ensatina eschscholtzii*

Abstract

Salamanders of the *Ensatina eschscholtzii* complex have featured prominently in evolutionary biology because they represent one of the few examples of a ring species. The notion of reproductive isolation between the geographically terminal forms of the ring (*E. e. eschscholtzii* and *E. e. klauberi*) in southern California is paramount to the ring species interpretation; previous work based on allozymes has shown that hybridization is infrequent or absent at the end of the ring. Here, we report on a fine-scale genetic analysis of a hybrid zone between the terminal forms. 335 individuals were genotyped at three diagnostic nuclear loci and one mitochondrial locus. Individual-based Bayesian methods for hybrid identification and classification revealed a higher frequency of hybridization than has been reported previously at other contact zones at the end of the ring. Pure parentals and F1s comprised the majority of the sample, with an apparent skew towards *eschscholtzii* alleles in backcrossed individuals. Although parental forms show differences in habitat and elevation across the contact zone, maximum-likelihood cline analysis using a model including variation in habitat-genotype associations did not significantly improve the fit of the clines to the data. Cline widths are concordant and narrow with respect to dispersal (widths of 718-799 m), but there is cytonuclear discordance, both in terms of introgression and the geographic position of mitochondrial versus nuclear clines. Nearly all hybrids possess mitochondrial DNA from only one parental type (*klauberi*) suggesting asymmetric isolation. Reproductive isolation is strong, but not complete, at this contact zone. Results presented here are consistent with a pattern of divergent selection stemming from local adaptation to alternative selection regimes that is promoting the evolution of reproductive isolation and maintaining lineage distinctiveness despite some gene flow, as is predicted for the terminal forms of a ring species.

Introduction

Hybrid zones have received special attention from evolutionary biologists because they provide a rare opportunity to study the factors that promote reproductive isolation in natural systems (Barton and Hewitt 1985; Harrison 1990; Mallet 2005). The *Ensatina eschscholtzii* ring species complex is an ideal system for studying species formation because it is composed of multiple lineages at various stages of the speciation process, including several zones of secondary contact and hybridization (Wake 1997; Pereira and Wake 2009). Ensatina salamanders inhabit mesic, forested environments in Pacific western North America from southwestern British Colombia to northern Baja California and inland along the Cascades and western slopes of the Sierra Nevada, as well as the Transverse and Peninsular ranges to the south (Fig. 1). In his detailed analysis of geographic variation and speciation in *Ensatina*, Stebbins (1949) hypothesized that the complex originated in northern California or southern Oregon, where an ancestral population divided into

two subpopulations that expanded southward around the arid Central Valley of California along two separate paths, one along the relatively low-elevation coast ranges, the other inland along the western slopes of the higher-elevation Sierra Nevada. Eventually, the two lineages came back into contact in southern California, creating a ring encircling the Central Valley. The coastal lineages evolved relatively uniform, unblotched color patterns (except one form, *picta*), while the inland, montane lineages evolved bold, blotched color patterns, which Stebbins thought represented different strategies for predator avoidance (crypsis via background matching in unblotched forms, crypsis via disruptive coloration in blotched forms, and mimicry in *xanthoptica*). Previously, multiple species were recognized, but Stebbins (1949) defined *Ensatina* as a Rassenkreis with seven subspecies (Fig. 1). Today, geographically adjacent forms intergrade in a chain of populations around the Central Valley, except in southern California where two distinct forms occur in sympatry with little or no hybridization. Decades of genetic work by D. B. Wake and colleagues using allozymes and mitochondrial DNA have uncovered remarkably high, geographically structured genetic diversity within *Ensatina*, suggesting the biogeographic history of *Ensatina* is complex, having featured periods of geographic isolation and multiple instances of secondary contact (Wake and Yanev 1986; Moritz et al. 1992; Jackman and Wake 1994; Kuchta et al. 2009a; Kuchta et al. 2009b; Pereira and Wake 2009). Overall, these studies are consistent with Stebbins' general biogeographic hypothesis in that genetic distances are higher between the geographically terminal populations than most adjacent populations (Moritz et al. 1992; Pereira and Wake 2009), and within subspecies there is a positive correlation between genetic and geographic distance (Jackman and Wake 1994).

Given that the extensive north-south oriented range of *Ensatina* encompasses substantial environmental heterogeneity, local adaptation of populations to alternative selection regimes is likely an important driver of differentiation within the group (Stebbins 1949). Environmental heterogeneity can play an important role in shaping hybrid zone structure. Cline theory posits that areas of differentiation along environmental gradients are maintained by a trade-off between selection and dispersal: local adaptation favors one parental form in one habitat, and the other in another (Haldane 1948; Fisher 1950; Endler 1973; Slatkin 1973; May et al. 1975; Endler 1977; Armsworth and Roughgarden 2008). But migrants dispersing into the neighboring habitat keep the populations mixed and prevent further local adaptation. A key assumption of this theory is that dispersal is random with respect to phenotype and the environment. However, it's much more likely that an individual's decision about whether and where to disperse depends on its fitness in a given environment, with different genotypes moving into preferred habitat patches where they are locally adapted (Armsworth and Roughgarden 2008; Edelaar et al. 2008). When genotype-habitat associations exist, simple clinal hybrid zone models may not effectively describe spatial variation in allele frequencies (Howard 1986; Harrison and Rand 1989; Bridle et al. 2001). Such associations may result from variation in the strength of selection in different habitats (Slatkin 1973), habitat preference (MacCallum et al. 1998), long-distance dispersal (Bridle et al. 2001), assortative mating (M'Gonigle and FitzJohn 2010), or some combination of these.

In this study, we present results of a fine-scale genetic analysis of a hybrid zone between the geographically terminal forms of the ring, the Monterey Ensatina (*Ensatina eschscholtzii eschscholtzii*) and the Large-blotched Ensatina (*Ensatina eschscholtzii klauberi*). Previous conclusions that these forms are reproductively isolated (or very nearly so) are paramount to the ring species interpretation, and to unresolved debates about species boundaries in this complex (Frost and Hillis 1990; Graybeal 1995; Highton 1998; Wake and Schneider 1998). Although

eschscholtzii and *klauberi* had not been found in sympatry when Stebbins first published his biogeographical hypothesis in 1949, it was later discovered that these forms meet in at least four geographically isolated contact zones in southern California (Stebbins 1957; Brown and Stebbins 1964; Brown 1974; Wake et al. 1986). Based on allozymes and color-pattern data, the frequency of hybridization varies among the contacts from rare hybridization to complete reproductive isolation in sympatry in the southernmost contact zone in the Cuyamaca Mountains (Wake et al. 1986; Wake et al. 1989). Although previous studies were sufficient to document geographic variation in hybridization frequency at the end of the ring (Wake et al. 1989), more detailed genetic analyses are needed to fully understand the nature and extent of reproductive isolation there. Here, we use Bayesian methods to identify and classify hybrids in combination with maximum-likelihood cline analysis to estimate cline shape and linkage disequilibrium, and from these parameters infer rates of dispersal and level of selection against hybrids. We test for habitat-genotype associations, and whether the incorporation of variation in habitat type and elevation improves cline fit (Bridle et al. 2001).

Materials and Methods

SAMPLING

The study site is located on Palomar Mountain, San Diego County, California (Fig. 1). Suitable habitat consists of mixed montane woodland and montane riparian forest dominated by Incense Cedar (*Calocedrus decurrens*), White Fir (*Abies concolor*), Black Oak (*Quercus kelloggi*), Coast Live Oak (*Quercus agrifolia*), Bigcone Douglas Fir (*Pseudotsuga macrocarpa*), and Canyon Live Oak (*Quercus chrysolepis*). The area sampled is approximately 3.5 km x 1.75 km. The distribution of sampled populations primarily follows a northeast-facing slope and boulder-filled creek adjacent to an open treeless meadow of unsuitable habitat (Fig. 2). Salamanders were captured using visual surveys of natural cover objects during the day and night driving. All individuals that were encountered were sampled, and latitude, longitude, and elevation (with error estimates <10m) were recorded at the point of capture using a GPS. 335 salamanders were sampled over a three-year period from January-April of 2005-7. Individuals were identified based on color pattern as *klauberi*, *eschscholtzii*, or hybrids (Fig. 1). Most individuals were sampled non-lethally by removing a piece of the tail tip (~5mm) to avoid collecting large numbers of individuals, and to allow for future long-term monitoring; a subset of individuals were euthanized and preserved as voucher specimens and deposited in the Museum of Vertebrate Zoology (MVZ), University of California, Berkeley. Tissue samples collected in the field were stored in 95% ethanol or propylene glycol and later frozen at –80ºC in the lab. Individuals that were not collected were photographed and marked using subcutaneous alphanumeric tags (Northwest Marine Technology, Shaw Island, WA) and returned to their point of capture within 24 hours.

MOLECULAR MARKERS

DNA was extracted from tissues (liver or tail-tip) using Qiagen DNeasy tissue kits following the manufacturer's protocol (Qiagen, Valencia, CA). Using PCR, we amplified three autosomal loci and one mitochondrial locus for all 335 individuals. PCRs consisted of 40 cycles of 94˚C for 30

s, Ta˚C for 45 s, and 72˚C for 60 s, with locus specific annealing temperatures (Table 1). PCR products were purified using ethanol following standard methods. The mitochondrial locus was sequenced in both directions; autosomal loci were sequenced in the forward direction only (Table 1). Purified templates were sequenced using dye-labeled dideoxy terminator cycle sequencing. DNA sequences were edited and aligned using Geneious Pro v4.7 (Drummond et al. 2009). We resolved haplotypes using PHASE (Stephens et al. 2001; Stephens and Donnelly 2003) and used SeqPHASE (Flot 2009) for converting between FASTA and PHASE files.

IDENTIFICATION AND CLASSIFICATION OF HYBRIDS

To estimate levels of admixture, I used Structure (Pritchard et al. 2000; Falush et al. 2003, 2007) and NewHybrids (Anderson and Thompson 2002; Anderson 2008) to analyze the multilocus sequence data. In the context of a two-population hybrid zone, Structure jointly assigns individuals probabilistically to the two parental populations. NewHybrids uses a different model, to compute the posterior probability that an individual belongs to one of six distinct genotype frequency classes (parentals, F1s, F2s, and backcrosses) that are possible following first- and second-generation matings between two species (Anderson and Thompson 2002; Anderson 2008). I used the linkage model in Structure and assumed two populations. Different gene fragments were treated as unlinked. I assumed sites within gene fragments were linked and that map distances between sites were proportional to the number of base pairs between sites (Falush et al. 2003). For Structure, I ran 100000 sweeps of five chains after a burn-in of 50000 sweeps and checked for convergence by comparing the estimated membership coefficient (Q) for each individual across the 5 runs. For NewHybrids, I ran 100000 sweeps of five chains started from overdispersed starting values after a burn-in period of 50000 sweeps following the software author's recommendation. Uniform priors were used for the mixing proportions and allele frequencies. To check for convergence, I visually inspected $P(z)$ values from the different runs which were then averaged across the 5 runs. Information about the population of origin of individuals was not used for either analysis. A threshold Q-value of 0.1 was used as a criterion in separating hybrids from pure parentals (i.e., an individual with a value between 0.1 and 0.9 is a hybrid, and any individual with a Q-value < 0.1 or > 0.9 is a pure parental) (Vähä and Primmer 2006). The same threshold was used to distinguish hybrids from pure parentals for the NewHybrids analysis, with posterior probability values summed across hybrid classes for an individual (Anderson and Thompson 2002; Anderson 2008).

TESTING FOR HABITAT-GENOTYPE ASSOCIATIONS

We tested whether hybrids and pure parentals show differences in elevation and broad-scale habitat type to see if a simple clinal hybrid zone model was appropriate for describing spatial variation in allele frequency. If significant genotype-habitat associations exist, the assumptions of a clinal model of hybrid zone structure may be violated (Harrison and Rand 1989; Bridle et al. 2001). We used elevation estimates taken at the point of capture for each individual and vegetation data from an unpublished floristic study of Palomar Mountain State Park (Lauri 2004). Individuals were classified as coming from one of two vegetation series, Mixed Montane Woodland (MMW) or Montane Woodland with *Pseudotsuga macrocarpa* (MWP; Lauri's [2004] Fig. 28). These series intergrade (Lauri 2004) and do not have sharply defined borders in and around the hybrid zone, but there is a discernible transition from MMW in the northwestern,

lower-elevation portion of the transect to MWP in the southeastern, higher-elevation portion of the transect (Fig. 2). Statistical analyses were performed using SPSS Statistics 17.0 (IBM).

CLINE ANALYSIS

Phased haplotypes were collapsed into a two-allele system (i.e., "*eschscholtzii*" and "*klauberi*" alleles) for cline analysis. All loci were diagnostic, allowing nearly all haplotypes to be unambiguously assigned to one parental form or the other. Two haplotypes representing four individuals for one of the loci (SLC8A3) could not be unambiguously assigned and were excluded from the cline analysis. We assumed that allele frequencies at individual loci did not change significantly over the three-year sampling period. Individuals were not pooled into discrete samples for cline fitting. A hybrid index summed across nuclear loci and scaled from 0 to 1 was calculated for each individual, expressed simply as the proportion of alleles derived from one of the two parental populations (in this case, *klauberi*) (Barton and Gale 1993). Sampling sites were collapsed onto a one-dimensional transect using the Pooled Adjacent Violators Algorithm (Brunk 1955), a method for finding the maximum-likelihood (ML) monotonic cline over a set of observations (Macholán et al. 2008). The advantage of the PAVA method is that it doesn't depend on a particular cline model; it assumes only a monotonic increase or decrease in allele frequencies across sampling sites (Macholán et al. 2008). Because the path of the contact front is assumed to be linear in the PAVA method, the best-fit axis of the transect orientation through sampled individuals was estimated using a one-dimensional cline fitting with a straight center line; we did not estimate the best-fit axis using more complex center lines comprising two or more segments (Bridle et al. 2001). Monotonic clines and their likelihood profiles were estimated using a routine written in Mathematica (Wolfram 1992) by S. J. E. Baird. The width of a cline is usually defined as the inverse of the maximum slope (Endler 1977). However, because it is not clear how best to estimate the maximum slope from monotonic clines, we estimated cline widths by fitting parametric sigmoid clines using Analyse v1.3 (Barton and Baird 1995). A sigmoid cline in allele frequencies (p) can be modeled by a tanh function of its width and center, such that

$$p = \left(1 + \tanh\left[2(x - c)/w\right]\right)/2 \quad (1)$$

where c is the cline center, w is the cline width and (x - c) is the distance from the cline center (Szymura and Barton 1986). We fit sigmoid clines and did not explore more complex stepped cline models because doing so only seems justified when there is sufficient sampling in the tails of the cline (Barton and Gale 1993). Parameter estimates are given as maximum-likelihood estimates, along with two-unit support limits analogous to 95% confidence intervals (Edwards 1992). Cline shape concordance among loci was assessed using the polynomial-fitting method of Szymura and Barton (1986) and the likelihood method described by Phillips et al. (2004). In Szymura and Barton's (1986) method, differences among loci were estimated by fitting a cubic polynomial relating the allele frequency at a particular locus (p_i) to the average frequency ($\bar{p}$) in the sample such that

$$p_i = \bar{p} + 2\bar{p}\bar{q}\left(\alpha + \left(\bar{p} - \bar{q}\right)\beta\right) \quad (2)$$

where α describes an increase in frequency of *klauberi* alleles above the average (i.e., a shift of the cline towards *eschscholtzii*'s range) and β describes an increase in frequency of *klauberi* alleles on the *klauberi* side and a decrease on the *eschscholtzii* side (i.e., a narrowing of the cline) (Szymura and Barton 1986). In units of cline width (defined as the inverse of the maximum slope of the cline (Endler 1977)), α is twice the shift in position, and β is the decrease in cline width below the average (Szymura and Barton 1986). If the clines for different loci are concordant, they would coincide and follow a diagonal line. (However, because these parameters do not distinguish differences in the degree of introgression from differences at the center, we fitted clines to each locus separately as described above). In the method described by Phillips et al. (2004), the likelihood surface of each locus is explored stepwise along axes for both center position (c) and width (w) while allowing other parameters to vary at each point (Phillips et al. 2004). Likelihood profiles (Hilborn and Mangel 1997) are then constructed for both c and w and summed over all loci, resulting in a log-likelihood profile for the ML shared center or width (Phillips et al. 2004). The shared ML estimate can then be compared to the sum of noncoincident profile ML estimates using a likelihood ratio test (Hilborn and Mangel 1997). Twice the difference in log likelihood ($G = 2\Delta LL$) between the two models under comparison is significant at level α if $G = 2\Delta LL > \chi^2_{df,\alpha}$ with the degrees of freedom equal to the difference in the number of parameters between the two models (Szymura and Barton 1986, 1991). We also tested whether habitat-genotype associations contribute to the observed spatial structure of the hybrid zone by comparing models incorporating variation in habitat and elevation to one assuming no habitat-genotype associations using a likelihood-ratio test (Bridle et al. 2001). Linkage disequilibrium (D) was calculated for the central region of the cline using a modified version of Hill's (1974) likelihood model implemented in ClineFit (Porter et al. 1997). Assuming that linkage disequilibrium is maintained by a balance between dispersal and selection, dispersal (σ, defined as the standard deviation of distances between birthplaces of parents and offspring) can be calculated from the estimate of cline width and linkage disequilibrium such that

$$\sigma = \sqrt{\frac{Drw^2}{1+r}} \tag{3}$$

where r is the harmonic mean recombination rate among loci under selection (assumed to be ~0.5 for unlinked loci (Szymura and Barton 1986)), and w is the cline width (Szymura and Barton 1991; Barton and Gale 1993). Given estimates of cline width (w) and rate of dispersal (σ) the effective strength of selection (s^*) on a locus at the center of the zone can be estimated following Barton and Gale (1993):

$$s^* = 8(\sigma / w)^2 \tag{4}$$

Results

HYBRID ZONE GENOTYPES

Consistent with initial identification of individuals based on color pattern (Fig. 2), there is a clear transition at all loci from *eschscholtzii* alleles in the northwest to *klauberi* alleles in the southeast. Classification of individuals as hybrids or pure parentals based on color pattern was consistent with the Structure and NewHybrids classifications, with 46 individuals classified as hybrids with high posterior probability (> 0.9) (Fig. 3A). 18 individuals with complete data were heterozygous at all three nuclear loci (F1s by definition) and another was heterozygous at two loci but missing data at the third locus. The majority of the remaining hybrids were classified as backcrosses with *eschscholtzii* (Fig. 3C). Of the 46 hybrid individuals, 43 (93%) possessed *klauberi* mtDNA (Fig. 3B). Of the three hybrids with *eschscholtzii* mtDNA, one was classified as an F1 ($pp > 0.99$) and two were classified as backcrosses with *eschscholtzii* ($pp > 0.98$ and $pp > 0.91$, respectively).

GENOTYPE-HABITAT ASSOCIATIONS

The two taxa are associated with different vegetation series in the contact zone (Chi-square test, $p < 0.001$; Fig. 4). *Eschscholtzii* are found almost exclusively in the MMW series along with >80% of the hybrids, while *klauberi* individuals are found at roughly equal frequencies in both the MMW and MWP series. *Eschscholtzii* occupies a broader, but lower elevational range (1200-1601 m; mean 1360 m) than *klauberi* (1297-1694 m; mean 1470 m) and hybrids (1342-1601 m; mean 1445 m) (One-way ANOVA, Tukey's HSD, $p < 0.001$; Fig. 5). Despite these associations, modification of the underling clines according to vegetation type (Bridle et al. 2001) gave no significant improvement in cline fit, and modification according to elevation provided only a slight improvement (results not shown).

TRANSECT ORIENTATION, CLINE SHAPE, AND CLINE CONCORDANCE

The best-fit axis of orientation according to the PAVA method equated to a heading of 207º (support limits 200º, 211º; Fig. 6). There was no strong evidence for different orientation among different loci (Fig. 7). Cline width and center estimates (along with support limits) along with the α and β components of Szymura and Barton's (1986) polynomial equation for assessing cline concordance across loci are provided in Table 2. Both mitochondrial and nuclear loci change over approximately the same distance (cline widths of 718-799 m). Plots of allele frequencies at each locus (p_i) against the average frequency ($\bar{p}$) were concordant for the three nuclear loci (Fig. 8B-D). The mitochondrial locus however, was not concordant with the nuclear loci (Fig. 8A), showing an increase in *klauberi* alleles above the average ($\alpha = 0.753$). This discordance represents a shift of the mitochondrial cline to the west by approximately 15% of the consensus cline width toward the range of *eschscholtzii*. Likelihood-ratio tests showed no significant difference in cline width when all four loci were considered separately versus when they were considered together.

DISPERSAL AND SELECTION ESTIMATES

Because cline widths (and centers) among nuclear loci were not statistically significantly different, dispersal and selection estimates are based on the multilocus nuclear cline width estimate (765 m). Dispersal (σ) estimated from the multilocus cline width was 124.6 (99.3 - 258) m/gen$^{1/2}$. Linkage disequilibrium (D) for the central region of the cline was 0.171 (0.148 - 0.250).

Effective selection against hybrid genotypes in the center of the cline (s^*) was estimated to be 21%.

Discussion

IDENTIFICATION AND CLASSIFICATION OF HYBRIDS

Initial identification of individuals as hybrids or pure parentals based on color pattern was consistent with genetic assignments. Eighteen of the 46 hybrid individuals were classified as F1s with high posterior probability (>0.99). The remainder of the sample is dominated by backcrosses with *eschscholtzii*. Most F2s and older generation backcrosses had lower posterior probabilities (Fig. 3). The accurate identification of hybrids based on multilocus genetic data depends on a number of factors, including level of divergence between parental forms, number of loci, proportion of hybrids in the sample, and number of generations of backcrossing (Boecklen and Howard 1997; Vähä and Primmer 2006). When there are fixed allelic differences between species, as few as four or five loci may be sufficient for distinguishing hybrids versus parentals (Boecklen and Howard 1997). Classifying individuals into distinct hybrid categories with confidence is a more difficult problem, however, especially when the number of generations of backcrossing is high (Anderson and Thompson 2002; Anderson 2008). The model implemented in NewHybrids assumes only two generations of interbreeding; in cases where the number of generations of potential interbreeding is greater than or equal to three, it becomes very difficult to classify individuals arising from many generations of backcrossing with any certainty, even with many diagnostic markers (Boecklen and Howard 1997; Anderson and Thompson 2002). Given that hybridization has been observed since the early 1960's (Brown 1974) and assuming a generation time of ~4 years (Stebbins 1954), it's reasonable to assume that more than three generations of interbreeding have taken place. Due to the limitations of the model and the difficult nature of the problem, classification of individuals as F2s or older generation backcrosses should be interpreted with some caution. The time since first contact between *eschscholtzii* and *klauberi* on Palomar Mountain is unknown, and any inferences about the age of the hybrid zone are purely speculative. However, it seems likely that *eschscholtzii* and *klauberi* first came into contact during the late Quaternary as a result of vegetation shifts associated with the climatic oscillations during the Pleistocene and Holocene, and the hybrid zone has reached a stable equilibrium.

Previous work on the contact zones between *eschscholtzii* and *klauberi* using allozymes and color pattern provides a useful background against which our results may be compared. Brown and Stebbins (1964) were the first to discover a hybrid zone between these taxa in southern California (Sawmill Canyon, San Bernardino Mountains). Based on color pattern, three of 15 individuals collected appeared to be hybrids. Brown (1974) later studied the Palomar Mountain hybrid zone, as well as the mid-ring contact in the Sierra Nevada between *xanthoptica* and *platensis*. He found that *eschscholtzii* and *klauberi* hybridized everywhere they met on Palomar, as did *xanthoptica* and *platensis* in the Sierra Nevada. Wake and Yanev (1986) and Wake et al. (1989) genotyped 37 individuals at 22 allozyme loci from six sites on Palomar Mountain over a much larger area than the one described here (~10x10km). A hybrid index was estimated, revealing a sample dominated by pure parentals but with several hybrids, including a single putative F1. Consistent with results presented here, most of the remaining hybrids were

dominated by *eschscholtzii* alleles. Of five individuals typed at the same loci from the Sawmill Canyon hybrid zone, two were classified as hybrids, both F1s. A sample of 13 individuals from the nearby San Jacinto Mountains also yielded two hybrids, neither of which were F1s. Finally, the largest sample of 51 individuals from the southernmost contact in the Cuyamaca Mountains where *eschscholtzii* is relatively rare did not yield any hybrids. Pereira and Wake (2009) reanalyzed the data from the four contact zones described in Wake et al. (1989) using NewHybrids (Anderson and Thompson 2002) and reached similar conclusions, finding that pure parentals dominated the samples (>80%), with most hybrids classified as F1s or first-generation backcrosses.

DISPERSAL AND SELECTION IN THE HYBRID ZONE

The indirect estimate of dispersal based on cline width was 124.6 (99.3 - 258) m/gen$^{1/2}$. Dispersal estimates for *eschscholtzii* or *klauberi* have never been made, but mark-recapture studies of other coastal and blotched forms of *Ensatina* have been conducted and provide useful direct estimates for comparison. Dispersal ability may differ between coastal, unblotched forms, and inland, blotched forms for a variety of reasons, particularly if these lineages have evolved different strategies for predator avoidance as first suggested by Stebbins (1949). Direct estimates of dispersal based on mark-recapture studies have been made for *xanthoptica* (females: max. 19.4 m, mean 5.3 m; males: max. 32.5 m, mean 10 m; (Stebbins 1954)), *oregonensis* (max. 42.6 m, mean 7.2 m (Lowe 2001)) and *platensis* (females: max. 60.6 m, mean 23.3; males: max. 120.4 m, mean 31.2 m (Staub et al. 1995)). It has been argued that direct estimates of dispersal are likely to represent underestimates because long-distance dispersal and colonization events are usually overlooked (Szymura and Barton 1986; Barton and Gale 1993). Notwithstanding assumptions of the method used to estimate dispersal, our indirect estimate of dispersal is on par with maximum movement distances that have been recorded for *Ensatina* (Staub et al. 1995).

In the face of dispersal by pure parentals into the hybrid zone at a rate of 124.6 m/gen$^{1/2}$, an effective level of selection $s^* = 21\%$ would be necessary to maintain a 765 m wide cline if selection acted against heterozygotes at a single locus. The particular markers themselves probably do not have a direct effect on fitness, but rather they are in linkage disequilibrium with other loci that are under selection (Szymura and Barton 1986). For comparison, selection against hybrids has been estimated at 32% for distinct lineages of lizards of the *Sceloporus grammicus* complex (Sites et al. 1995; Marshall and Sites 2001) and 17-22% for the toads *Bombina bombina* and *B. variegata* (Szymura and Barton 1986, 1991). Given such a high selection coefficient, introgression of neutral alleles is unlikely, although favorable mutations could move easily if prezygotic isolation is weak or absent (Pialek and Barton 1997; Barton 2000).

COMPARISON TO THE *PLATENSIS-XANTHOPTICA* HYBRID ZONE

A rare opportunity to compare our results to those from another secondary contact between morphologically and genetically distinct coastal and inland lineages of *Ensatina* is provided by the "mid-ring" hybrid zone between *xanthoptica* and *platensis* in the foothills of the central Sierra Nevada (Brown 1974; Wake et al. 1989; Alexandrino et al. 2005). These morphological analogs of *eschscholtzii* and *klauberi*, respectively, came into contact at some point during the Pleistocene or Holocene when conditions were more mesic, allowing *xanthoptica* to cross the Central Valley and invade the Sierras from the San Francisco Bay area (Stebbins 1949).

Alexandrino et al. (2005) used the same mathematical cline model described here to analyze eight allozyme loci, mitochondrial DNA, and color pattern at two different contact zones between these taxa in the Sierra Nevada. In contrast to this study, Alexandrino et al. (2005) found few if any F1s. They suggested that reduced opportunities for heterospecific encounters due to habitat preference and/or stronger selection against F1s compared to later generation hybrids could explain this pattern. Despite cline width estimates comparable to those observed here (730-2000 m), Alexandrino et al. (2005) inferred much stronger selection against hybrids (46-75%). As noted by Alexandrino et al. (2005), the abundance of hybrids in the contact zone seems at odds with strong selection against them. Levels of divergence are higher between *eschscholtzii* and *klauberi* (Nei's D = 0.6) compared to *xanthoptica* and *platensis* (Nei's D = 0.4) based on allozymes (Pereira, unpubl.), suggesting that reproductive isolation should be stronger between *eschscholtzii* and *klauberi*.

CLINE CONCORDANCE AND CYTONUCLEAR DISEQUILIBRIUM

The narrow width, shape, and overall concordance of clines for presumably unlinked loci suggest that, like many other animal hybrid zones that have been studied to date, this hybrid zone is maintained by a balance between dispersal by parental forms into the center of the hybrid zone and selection against hybrids (a "tension zone" (Barton and Hewitt 1985)). The discordance between the mitochondrial and nuclear loci, both in terms of introgression and the geographic position of clines, can provide insight into patterns of mating and mechanisms of selection in the hybrid zone (Arnold 1993). Given that nearly all of the hybrids possess *klauberi* mtDNA, either: 1) hybridization between *eschscholtzii* and *klauberi* is (mostly) unidirectional, with F1 hybrids formed from female *klauberi* mating with male *eschscholtzii* (but not vice versa, implying asymmetric prezygotic isolation), perhaps due to differences in the discriminatory power between females of the two species (Ehrman and Wasserman 1987; Kaneshiro and Giddings 1987), or 2) hybridization is reciprocal, but offspring resulting from female *eschscholtzii* mating with male *klauberi* (or male hybrids) are inviable or sterile (implying asymmetric postzygotic isolation) because one cross may produce fertile offspring, while the reciprocal cross is inviable (Harrison 1983; Hoskin et al. 2005; Turelli and Moyle 2007). The presence of three hybrid individuals with *eschscholtzii* mtDNA suggests matings between female *eschscholtzii* and male *klauberi* (or male hybrids) are possible, but appear to be rare, or that the offspring usually do not survive. Like many animals, female *Ensatina* are the choosier sex because they invest relatively more than males in reproduction, and are courted by promiscuous males, both of their own and closely related species (Andersson 1994; Coyne and Orr 2004). If there are costs associated with heterospecific matings (e.g., because hybrids are less fit, or, because these matings produce fewer offspring), there may be strong selection acting on female mating preferences toward increased ability to discriminate between conspecific and heterospecific males (McPeek and Gavrilets 2006). This in turn could drive indirect selection on males to track female mating preferences (Lande 1981).

Hybrids often possess the mitochondrial DNA of only one of two parental species in nature (Avise and Saunders 1984; Randler 2002; Lushai et al. 2005), a pattern predicted in species with female-choice when females of a rare species are unable to find conspecific males because they are scarce, and eventually accept matings with heterospecific males of a more common species (Wirtz 1999). However, males of the two species are equally abundant in zones of overlap at Palomar Mountain, suggesting that factors other than rarity of conspecific *klauberi*

males are responsible for the disproportionate percentage of hybrids with *klauberi* mtDNA. Because speciation is ultimately a by-product of local mating interactions, future work incorporating information about patterns of mating and gamete utilization will be critical for understanding reproductive isolating barriers and the role of selection in generating and maintaining species boundaries at the end of the ring (Marshall et al. 2002; McPeek and Gavrilets 2006).

Acknowledgements

We thank Craig Moritz, Jimmy A. McGuire, David B. Wake, and George K. Roderick for comments on an earlier draft of this manuscript. We thank Eric Anderson and Stuart J. E. Baird for guidance with analyses. We thank R. K. Lauri for providing an electronic copy of his unpublished thesis and vegetation map of Palomar Mountain State Park and Michelle Koo for assistance digitizing the map. We thank the ESRI Conservation Program for providing ArcGIS software for our use. For help in the field, we thank Mike Anguiano, Rayna Bell, Chuck Brown, Susan Cameron, Jessica Castillo, Becky Chong, Erin Conlisk, Brandon Endo, Matt Fujita, Emilio Gabbai-Saldate, Zach Hanna, Kory Heiken, Jasmine Junge, Megan Lahti, Ben Lowe, Matt McElroy, Craig Moritz, Greg Pauly, Ricardo Pereira, Tod Reeder, Mark Roll, Sean Rovito, Kevin Rowe, Frank Santana, Sean Schoville, Sonal Singhal, Tate Tunstall, and John Wiens. We thank Kim and Donna Rosier, Bill Stephenson, and the rest of the staff of the Palomar Mountain Christian Conference Center for their hospitality and access to their property. Permission to work in Palomar Mountain State Park was granted by the State of California Department of Parks and Recreation; we thank Mark Jorgensen, Nedra Martinez, and Jeff Lee for assistance and access to the park. This work was conducted under a scientific collecting permit issued by the California Department of Fish and Game (SC-007654). We thank Kevin Fleming and Art Fong (CDF&G) for assistance with permitting. Funding for this work was provided by Sigma Xi, the National Science Foundation (Doctoral Dissertation Improvement Grant DEB-0909821 to TJD and NSF DEB-0641078 to CM), the University of California Department of Integrative Biology, and the Museum of Vertebrate Zoology Martens and Louise Kellogg funds.

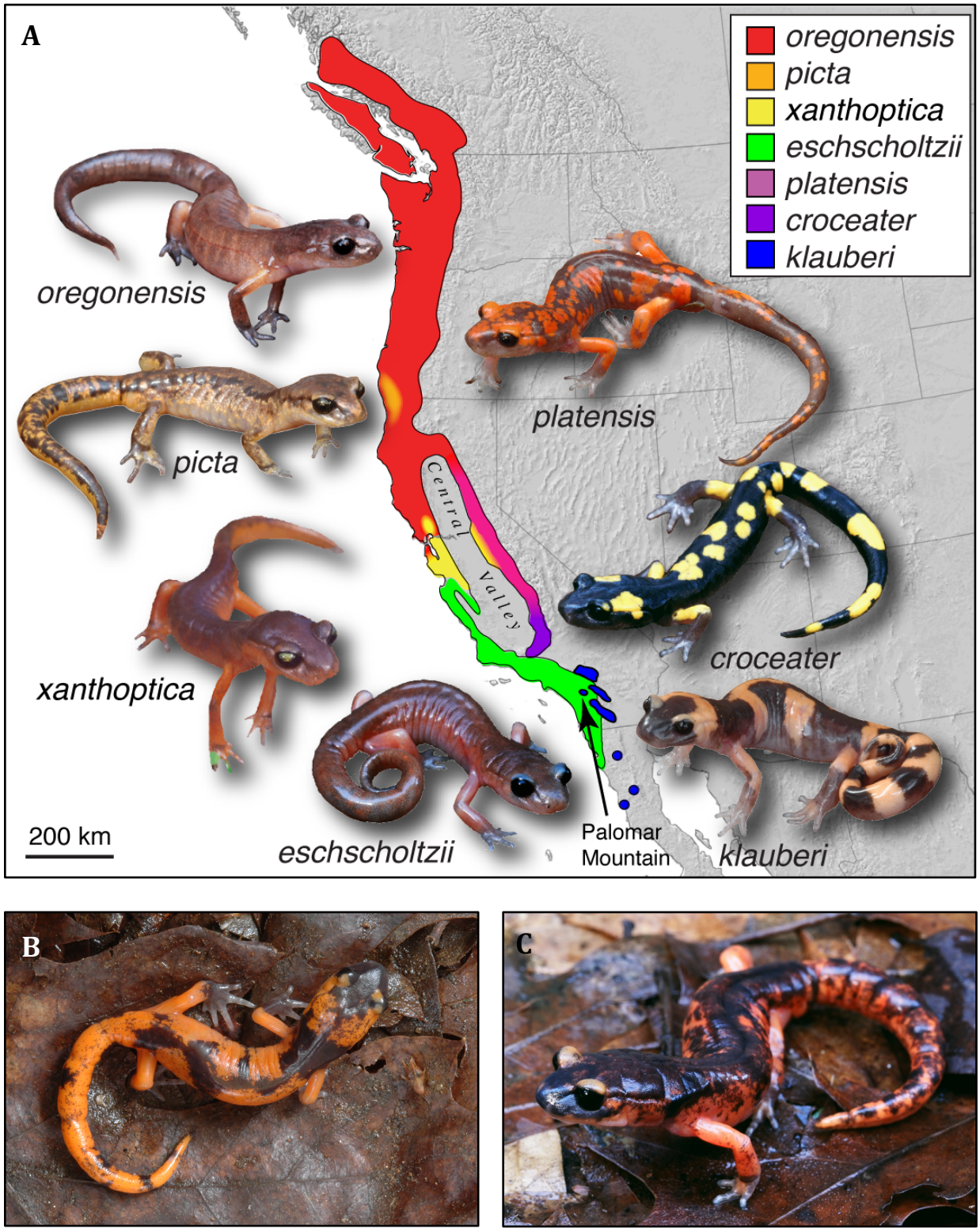

Figure 1. A) Range of the *Ensatina eschscholtzii* complex showing the location of the Palomar Mountain hybrid zone between *eschscholtzii* and *klauberi* described here; B-C) hybrid individuals showing intermediate color pattern.

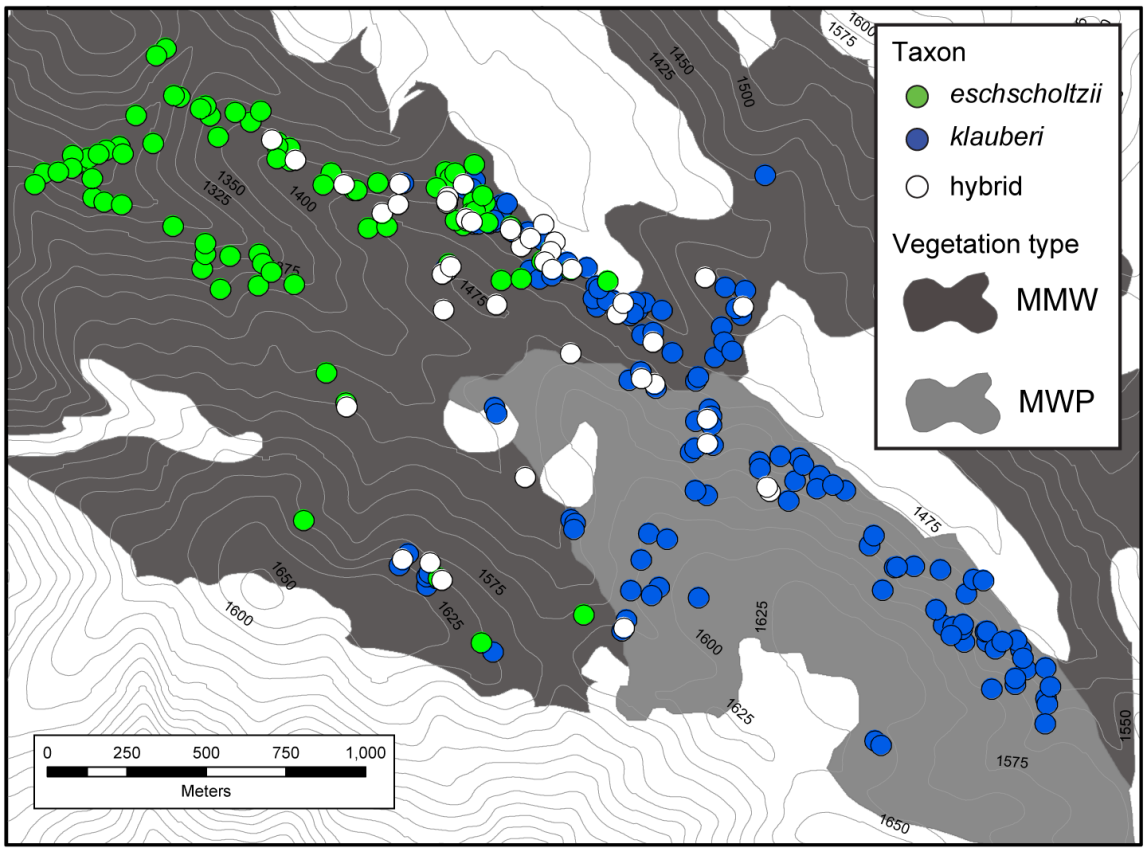

Figure 2. Distribution of *klauberi*, *eschscholtzii*, and hybrid individuals (*N*=335) across the contact zone (based on phenotype and genotype). Suitable habitat is represented by gray polygons, where MMW = Mixed Montane Woodland and MMP = Mixed Woodland with Bigcone Douglas fir (*Pseudotsuga macrocarpa*). Vegetation map modified from Lauri (2004).

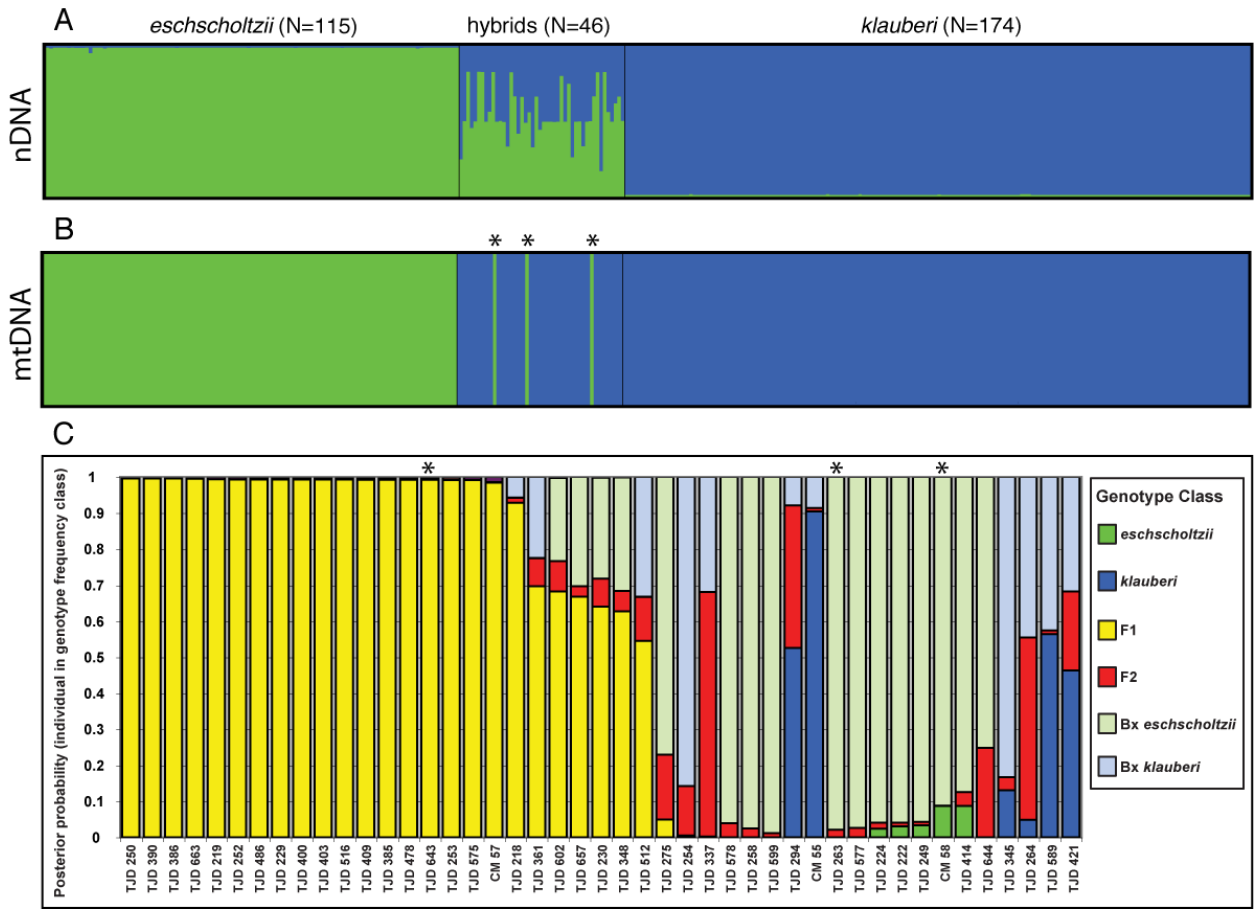

Figure 3. Identification and classification of hybrids. A) Structure results showing proportion of membership from each parental population for all 335 individuals; B) MtDNA haplotypes for all individuals; C) NewHybrids classification of the 46 hybrids classified by genotype frequency class. The three individuals with asterisks in (B) and (C) represent hybrids with *eschscholtzii* mtDNA.

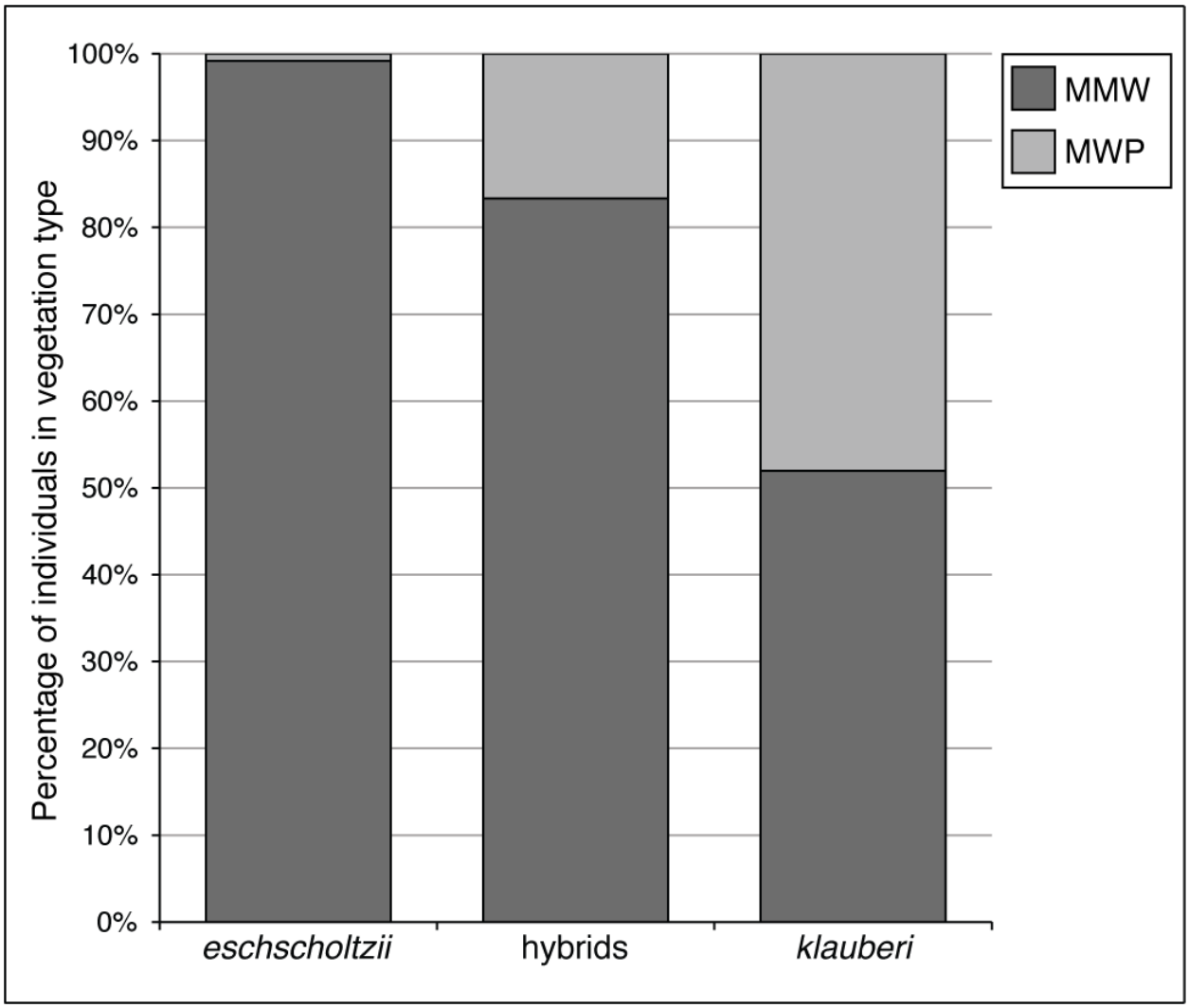

Figure 4. Differences in proportion of habitat type occupied by parentals and hybrids (Chi-square test, $p<0.001$).

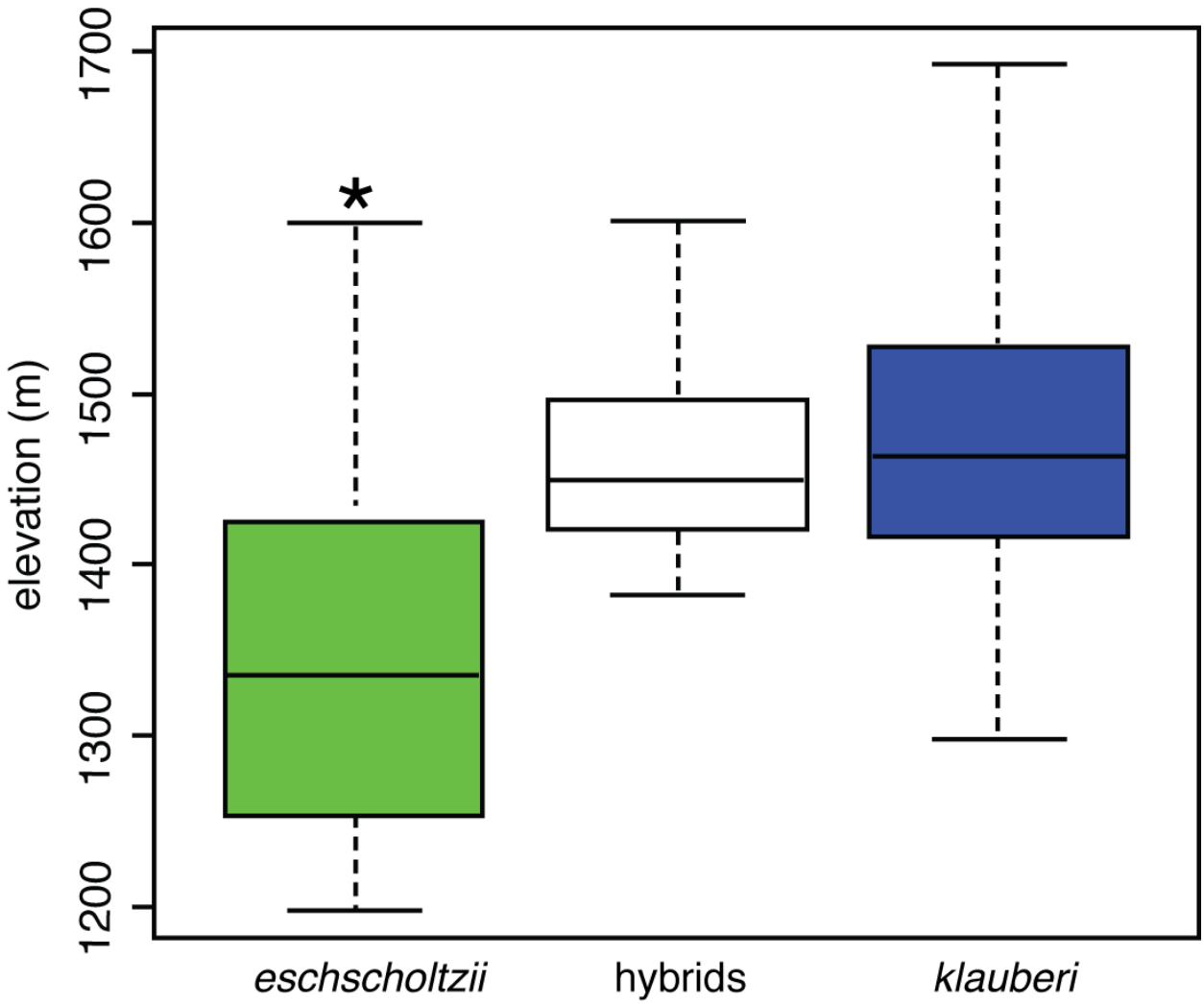

Figure 5. Distribution of parentals and hybrids by elevation. Box plots show the minimum, maximum, mean, and first and third quartiles. Sampled *eschscholtii* individuals in the hybrid zone occupy a statistically significant lower elevation range (Tukey's HSD test, $p<0.001$).

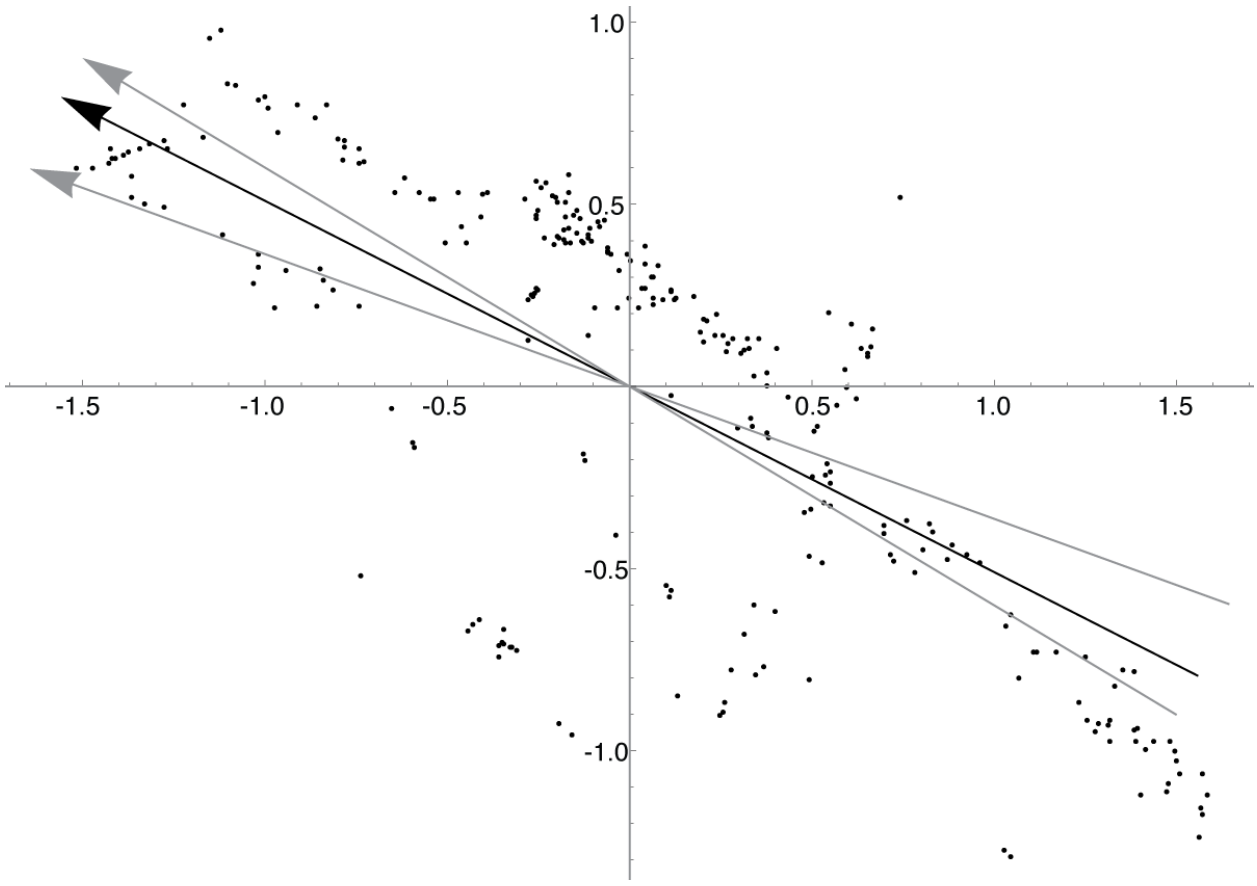

Figure 6. Orientation of change in allele frequencies in two-dimensional space estimated by the PAVA algorithm. Dots represent sampled individuals under a gnomonic projection around the mean latitude/longitude. Units are kilometers. The black arrow indicates the best-fit orientation axis, equating to a heading of 207°. Gray arrows represent lower and upper support limits (200°, 211°).

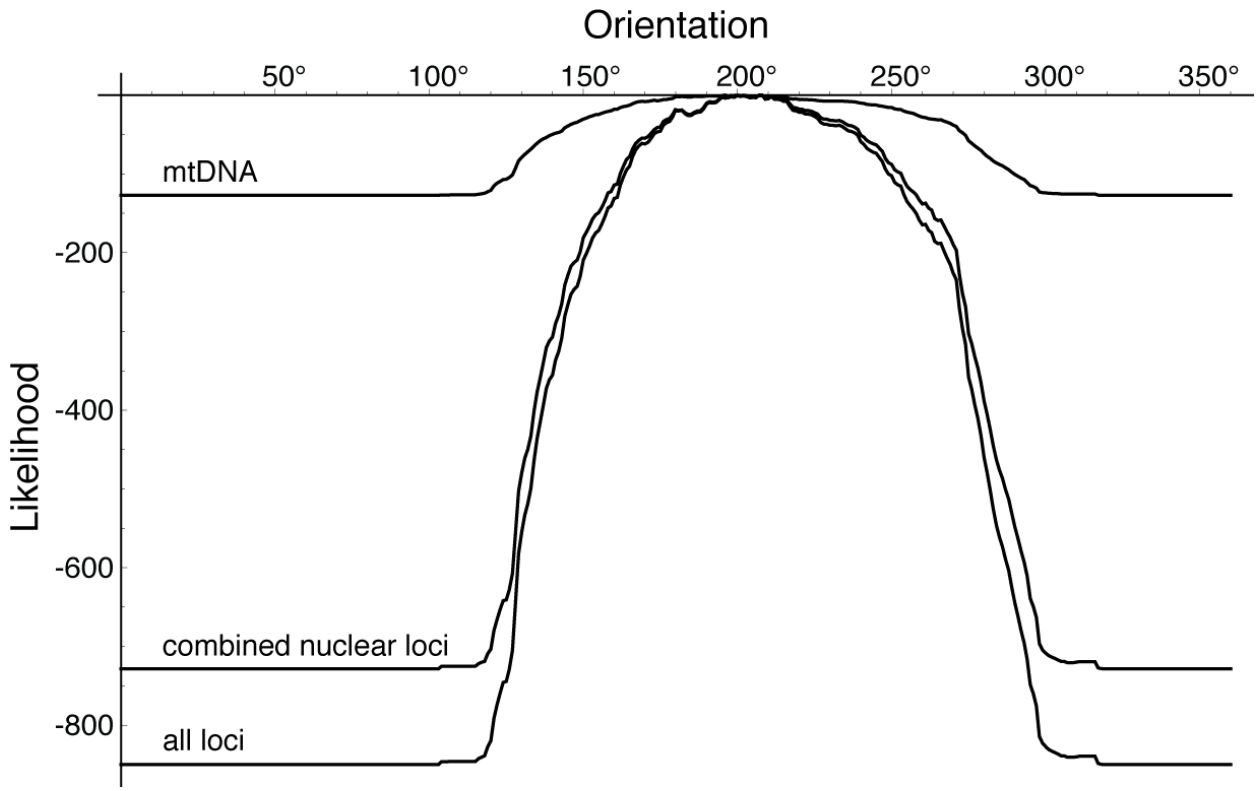

Figure 7. Support curves for the PAVA analysis of orientation of change in allele frequencies in two-dimensional space for the mitochondrial and nuclear loci.

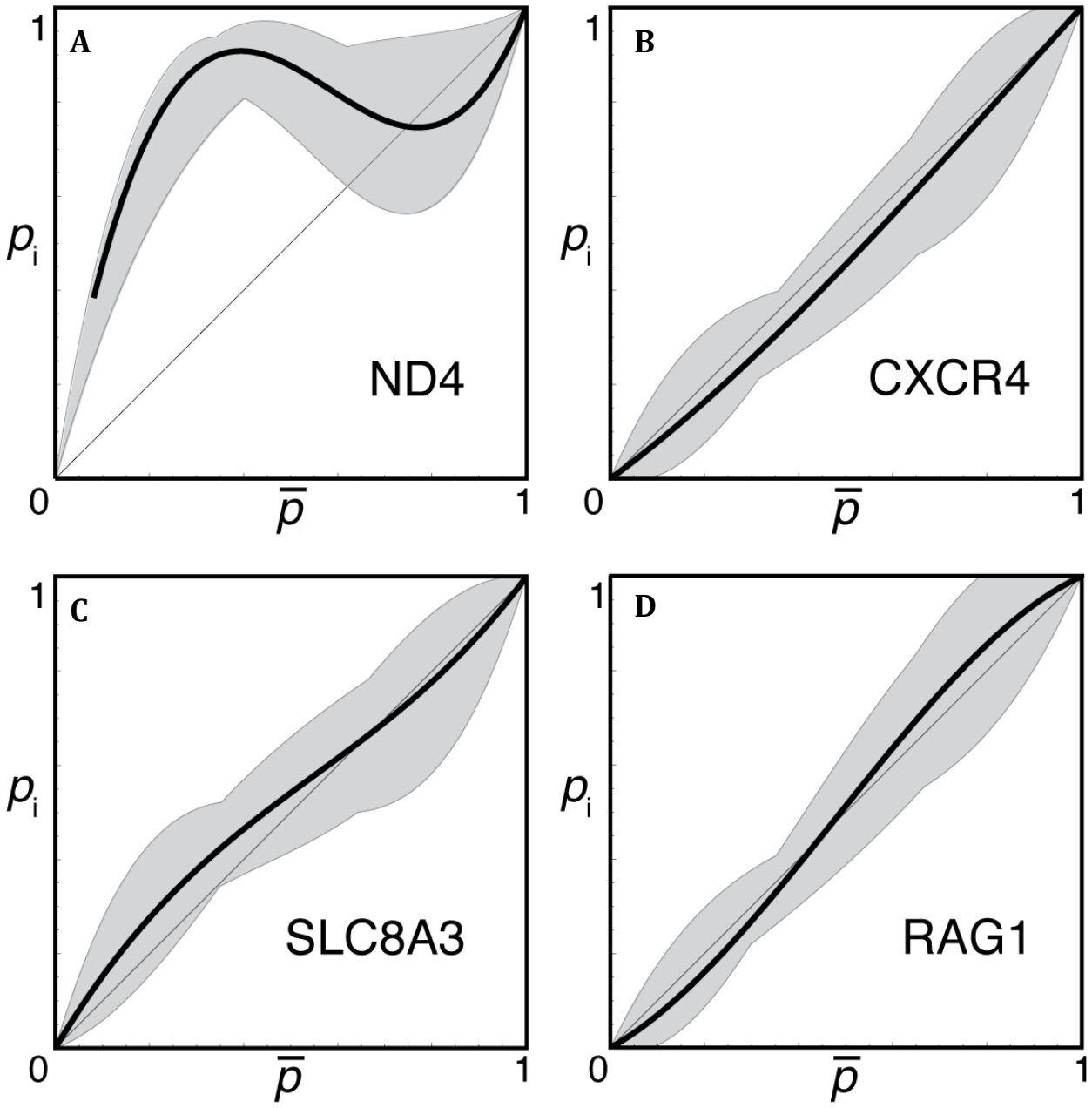

Figure 8. Comparison of mitochondrial (A) and nuclear (B-D) cline concordance. Plots show an individual's allele frequency at each locus (p_i) plotted against the average frequency in the total sample ($\bar{p}$) according to a cubic polynomial equation (see text). The nuclear clines are concordant with each other, while the mitochondrial cline is not.

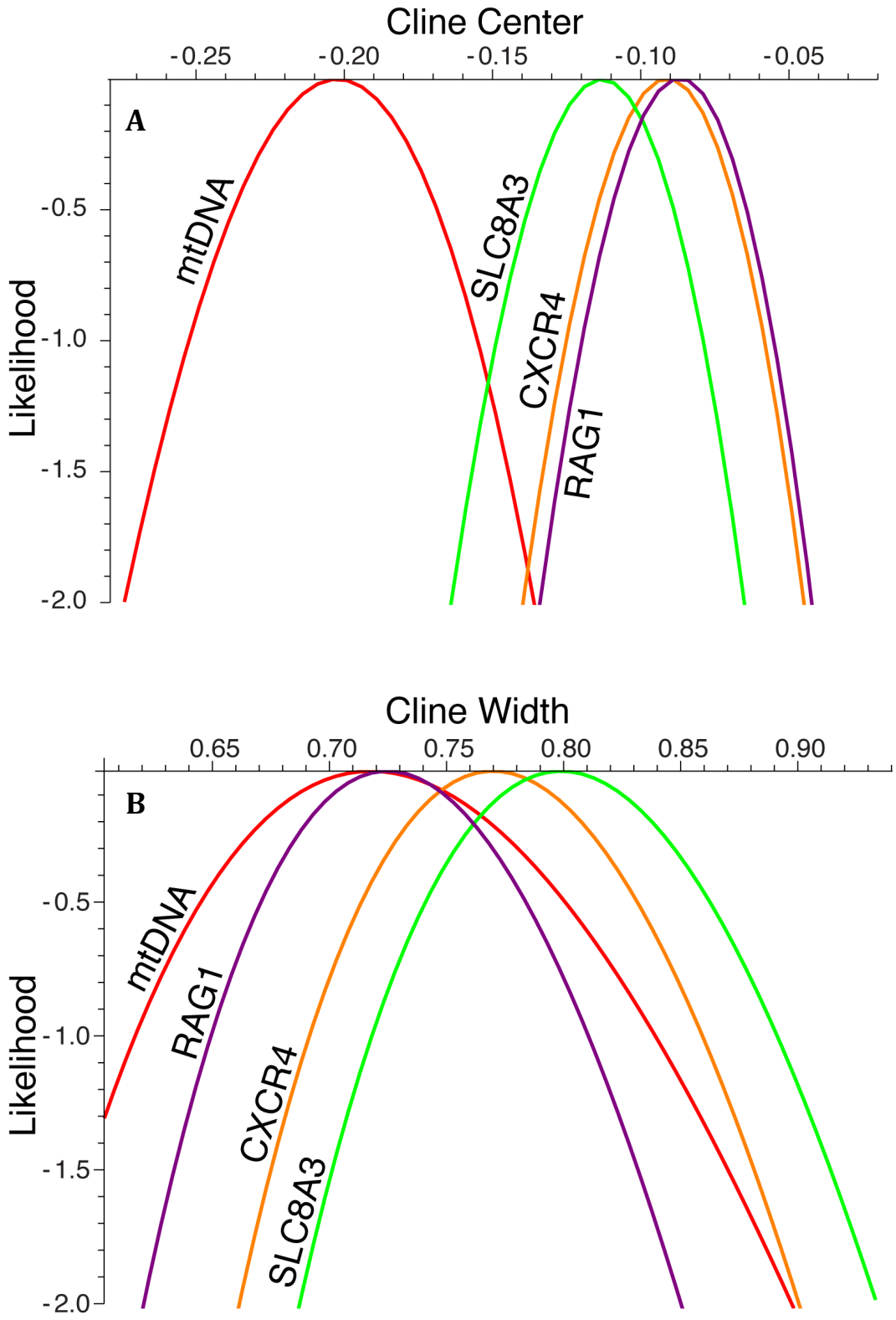

Figure 9. Maximum-likelihood profiles for the centers (A) and widths (B) of clines. Cline centers for the nuclear loci are concordant with each other, while the mitochondrial cline is shifted significantly to the west (A). There is no significant difference in cline widths (B).

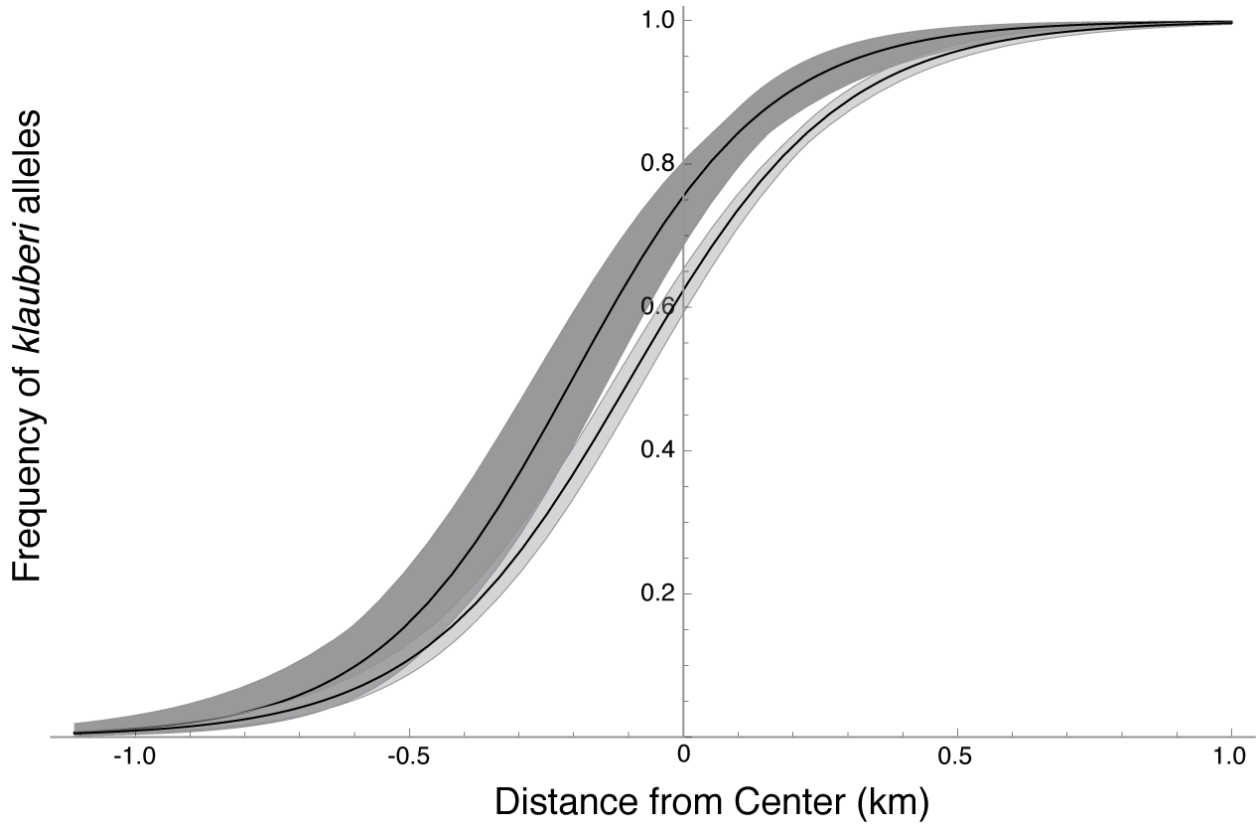

Figure 10. Mitochondrial (dark gray) versus consensus nuclear clines (light gray) along with support envelopes.

Table 1. PCR annealing temperatures, sequence length, and primer sequences.

Locus	Annealing temp.	Length/ # variable sites (bp)	PCR primers[1]
ND4	48°C	797/87	ND4
			LEU
CXCR4	55°C	116/4	TGGTCTGTGGATGCTGTCAT
			TGCAGTAGCAGATCAAGATGA
SLC8A3	58°C	342/6	CATTCGGGTCTGGAATGAAA
			ACACCACCATCCCCTCTGTA
RAG1	56°C	359/7	ACAGGATATGATGARAAGCTTGT
			TTRGATGTGTAGAGCCAGTGGTGYTT

[1]ND4 and LEU from Arévalo (1994); CXCR4, SLC8A3, RAG1 modified from Roelants et al. (2007).

Table 2. Maximum-likelihood estimates (with lower, upper support limits) of cline centers and widths along with the α and β components of the polynomial equation for assessing cline concordance. Center and width units are in kilometers and centers are measured with the gnomonic projection focus at zero. α describes an increase in frequency of *klauberi* alleles above the average (i.e., a shift of the cline towards *eschscholtzii*'s range) and β describes an increase in frequency of *klauberi* alleles on the *klauberi* side and a decrease on the *eschscholtzii* side (i.e., a narrowing of the cline).

Locus	Center	Width (km)	α	β
All loci	- 0.112 (- 0.138, - 0.087)	0.764 (0.704, 0.830)	--	--
ND4	- 0.202 (- 0.274, - 0.136)	0.718 (0.580, 0.898)	0.753	- 1.522
All nuclear loci	- 0.097 (- 0.125, - 0.070)	0.765 (0.700, 0.838)	--	--
CXCR4	- 0.091 (- 0.140, - 0.045)	0.770 (0.661, 0.901)	- 0.091	0.040
SLC8A3	- 0.113 (- 0.164, - 0.065)	0.799 (0.687, 0.934)	0.084	- 0.258
RAG1	- 0.087 (- 0.134, - 0.042)	0.725 (0.620, 0.851)	0.026	0.248

Chapter 2

Phylogeography of the Large-blotched Ensatina Salamander: Effects of Late Quaternary Climate Change on Population Divergence and Connectivity in the Peninsular Range Region of the California Floristic Province

Abstract

Late Quaternary climate change has had a profound impact on the distribution of species, genetic, and community diversity worldwide. Understanding the biotic effects of these climatic shifts remains a fundamental goal in evolutionary biology. The emerging field of ecological niche modeling provides a means of quantifying the spatial distribution of suitable environmental conditions for a species and can be used to model species distributions during periods of climatic extremes in the recent past to make spatially explicit demographic predictions that can then be tested using phylogeographic data. Here, we use species distribution modeling from the present and two different historical time periods during the late Quaternary (Last Glacial Maximum, 21 ka and mid-Holocene, 6 ka) to make predictions about current patterns of genetic structure in the Large-blotched Ensatina (*Ensatina eschscholtzii klauberi*), a Plethodontid salamander endemic to middle and high elevation sky-island conifer forest in the Eastern Transverse and Peninsular ranges of southern California and northern Baja California. Intersecting species distribution models across time periods predicts one large and several smaller stable refugia. Using allozymes, microsatellites, and mtDNA, we tested for phylogeographic structure overall, and asked specifically whether there was significant structure among refugia reflecting long-term isolation in these areas. Clustering analyses based on allozymes recovered a distinct northern and southern group separated by a major fault zone. Pairwise comparisons of genetic distances between populations were significant, but levels of differentiation were relatively low overall. Geographic structuring of variation by refugia was not supported. Results suggest that populations in putative refugia have not been isolated for very long, or that recurrent patterns of gene flow in response to cyclical environmental changes and habitat shifts may have masked any earlier periods of divergence in allopatry.

Introduction

Global climate has fluctuated greatly in the last 3 million years, and has had a profound impact on the current distribution of genetic, species, and community diversity (Hewitt 1996, 2000, 2004). In particular, the Last Glacial Maximum (LGM) 18,0000-21,000 years ago caused major shifts and fragmentation of species' ranges and their habitats. Understanding the biotic consequences of Pleistocene range shifts and fragmentation remains a fundamental goal in historical biogeography and evolutionary biology. A relatively new approach integrating the emerging field of ecological niche modeling with phylogeography has provided novel insight into lineage formation by comparing species paleodistributions with current phylogeographic patterns (Hugall et al. 2002; Graham et al. 2006). An extension of this approach uses paleodistribution models to predict stable refugia — areas where a species is predicted to have

persisted throughout climatic fluctuations — thereby forming spatially explicit hypotheses about historical demography which can then be tested using molecular data (Carstens and Richards 2007; Knowles et al. 2007; Richards et al. 2007; Waltari et al. 2007; Peterson and Nyári 2008; Carnaval et al. 2009; Bell et al. 2010).

Here, I explore the phylogeography of the Large-blotched Ensatina (*Ensatina eschscholtzii klauberi*). This direct-developing (i.e., no aquatic larval stage) lungless plethodontid salamander is endemic to the eastern Transverse Range (San Bernardino Mountains) and the Peninsular Range of southern California and northern Baja California where it inhabits geographically isolated stands of conifer and mixed conifer forest at middle and high elevations (Fig. 1) (Mahrdt et al. 1998; Stebbins 2003; Heim et al. 2005). The isolation of the fragmented "sky-islands" inhabited by this species is evidenced by the exceptional richness, rarity, and endemism of the Peninsular Range regional biota (Stephenson and Calcarone 1999). Southern California's forests have experienced major latitudinal and altitudinal shifts since the Last Glacial Maximum (LGM; 21-16 ka) (Minnich 2007). During the Pleistocene, cooler, wetter conditions allowed the expansion of forest into areas formerly occupied by desert, and formed pluvial lakes in lowland areas (Van Devender 1990). The transition from the Pleistocene to the modern Holocene climate around 11 ka marked a major shift to a warmer climate along with deglaciation (Minnich 2007). The mid-Holocene in California (5-3.5 ka) was a period of greater precipitation, expansion of Sierra Nevada glaciers, and formation of small pluvial lakes at lower elevations (Minnich 2007).

Because terrestrial salamanders are intimately linked to the environment physiologically (Feder 1983), they respond to even minute changes in microclimate or community succession (Davic and Welsh 2004) and as such, are excellent indicators of environmental change. If populations experienced long-term isolation in separate refugia during climatic oscillations, we might expect to see a pattern of strong genetic divergence among refugia. Conversely, if cyclical fluctuations have permitted recurrent population connectivity and gene flow, genetic divergence among populations may be low. Here, we use species distribution modeling for the present and potential distributions at two different time periods representing different climatic extremes during the late Quaternary (21 ka, and 6 ka) to make predictions about the distribution, genetic diversity, and population connectivity of the Large-blotched Ensatina. Allozymes, microsatellites, and mitochondrial DNA sequence data are then used to validate the model and test predictions using a suite of population and landscape genetic methods.

Methods

SPECIES DISTRIBUTION MODELING

Species distribution models were constructed using 116 unique locality records from the Museum of Vertebrate Zoology (MVZ), the San Diego Natural History Museum (SDNHM), and the Autonomous University of Baja California (UABC) (Fig. 1). I modeled the species' distribution for current conditions and at two different time periods in the recent past: 1) the Last Glacial Maximum (LGM) a cooler, wetter period ca. 21 ka; and 2) a warmer, wetter period during the Holocene (HOL) ca. 6 ka (Fig. 2A-C). Current climatic conditions were estimated from WorldClim (Hijmans et al. 2005), a database consisting of monthly means for temperature and precipitation for 1950-2000 at 1km spatial resolution. Paleoclimate conditions were

generated from downscaled estimates from the ECHAM3 model (Deutsches Klimarechenzentrum Modellbetreuungsgruppe 1992) available from WorldClim. Distribution models were constructed using the Maximum Entropy algorithm implemented in Maxent (Phillips et al. 2006). In order to obtain a biologically meaningful set of uncorrelated variables (Austin 2002), we removed those that were significantly correlated with other variables, resulting in a total of 10 variables (annual mean temperature, temperature seasonality [coefficient of variation across months], mean temperature of the wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, annual precipitation, precipitation seasonality, precipitation of driest quarter, precipitation of warmest quarter, and precipitation of coldest quarter). The model prediction was restricted to southern California (Imperial, Kern, Los Angeles, Orange, Riverside, San Bernardino, San Diego, and Ventura counties) and northern Baja California (Ensenada, Mexicali, Tecate, and Tijuana municipalities). I generated 5,000 pseudoabsence points from a background area containing the Transverse Ranges, Peninsular Ranges, and the coast to south of San Quintín, removing those points that were within 1 km of an actual occurrence record. To convert the Maxent results to a map indicating species presence and absence, we used the equal sensitivity and specificity threshold. The presence-absence maps for the three time periods were summed to estimate range stability through time. Stability values ranged from 0 (never present) to 3 (present during all three time periods) (Fig. 2D).

SAMPLING

Three molecular datasets were used to investigate phylogeographic patterns. The first is a dataset consisting of 23 allozyme loci scored from 97 individuals from California (N = 3-28 per locality; Fig. 3A; Appendix 1). The other two datasets consist of 70 (different) individuals sequenced for one mitochondrial gene and genotyped at 10 microsatellite loci from most of the same populations sampled for allozymes, but lacking the Sawmill Canyon population and with additional populations from Hot Springs Mountain and three populations in Baja California (N = 1-18 per locality; Fig. 3B; Appendix 2). Population names used throughout that are shared by both datasets do not necessarily correspond to the exact same locality because of sampling differences. GPS coordinates, with error estimates (< 10m), were taken at the point of capture for each individual or georeferenced from locality information (Wieczorek et al. 2004). Individuals from protected areas and well-sampled localities were sampled non-lethally by removing a piece of the tail tip (~5mm) and released; specimens from new localities were euthanized and preserved as voucher specimens following standard protocols. Tissue samples collected in the field (tail-tips) were stored in 95% ethanol or propylene glycol and later frozen at –80ºC in the lab.

MOLECULAR METHODS

DNA was extracted from tissues (liver or tail-tip) using Qiagen DNeasy tissue kits following the manufacturer's protocol (Qiagen, Valencia, CA). An ~800 base pair fragment of mitochondrial DNA (mtDNA) comprised of the ND4 gene and flanking tRNAs (His, Ser, Leu) was amplified via the polymerase chain reaction (PCR) using the primers ND4 and LEU (Arévalo et al. 1994) (GenBank Accession numbers xxxx-xxxx; Appendix X). PCRs consisted of 35 cycles of 94˚C for 1 min., 48˚C for 2 min., and 72˚C for 3 min. PCR products were purified using sodium acetate. Purified templates were sequenced using dye-labeled dideoxy terminator cycle

sequencing on an ABI 3730 automated DNA sequencer. DNA sequences were edited and aligned using Geneious Pro v4.8.5 (Drummond et al. 2009). Microsatellite loci were amplified and genotyped following Devitt et al. (2009) on an ABI 3730 DNA analyzer (Applied Biosystems, Inc.). Alleles were scored by hand using GeneMapper 4.0 (ABI). I visualized the scoring of alleles and checked for genotyping errors using MsatAllele_1.0 (Alberto 2009).

DESCRIPTIVE STATISTICS

For the microsatellites and allozymes, I calculated the observed and expected number of heterozygotes using Levene's (1949) correction. An exact test was used to test for Hardy-Weinberg equilibrium (HWE). F_{IS} was estimated following the method of Weir and Cockerham (1984). Significance was assessed using Markov chain Monte Carlo (MCMC) simulation (Guo and Thompson 1992) with default values for the Markov chain (100 batches with 1,000 iterations per batch and 1,000 dememorization steps) implemented in GenePop 4.0 (Raymond and Rousset 1995; Rousset 2008). Four of the allozyme loci were monomorphic across all populations and were excluded from subsequent analyses. The allozyme dataset was almost complete. A significant proportion of missing values in the microsatellite dataset were present however, due to population-specific allelic dropout (suggesting mutations in primer-binding regions), despite cross-specific PCR amplification (Devitt et al. 2009). Maximum-likelihood estimation of null allele frequency for the microsatellites was performed using the expectation-maximization (EM) algorithm (Dempster et al. 1977) implemented in GenePop 4.0 (Rousset 2008).

ESTIMATING THE NUMBER OF GENETIC CLUSTERS (K)

The number of genetic clusters (K) was estimated for the allozyme dataset using Structure 2.3.3 (Pritchard et al. 2000), Structurama 1.0 (Pella and Masuda 2006; Huelsenbeck and Andolfatto 2007) and Genodive 2.0b17 (Meirmans and Van Tienderen 2004). We estimated K in Structure using the admixture model assuming correlated allele frequencies between populations (the "F-model"), the best model for detecting structure between closely related populations (Falush et al. 2003). We explored values of K ranging from 1-6. Ten runs were conducted for each value of K, with each run consisting of 100000 sweeps after a burn-in of 50000 sweeps. We chose the best K following the method of Evanno et al. (2005). In Structurama, K was modeled as a random variable following a hierarchical Dirichlet process prior (Pella and Masuda 2006). We ran four chains of 100000 sweeps each sampled every 25 generations and discarded the first 1,000 samples as burn-in. Results are expressed as the mean population partition, the partitioning of individuals to populations that minimizes the squared distance to the partitions sampled by the MCMC algorithm (Huelsenbeck and Andolfatto 2007). We also used a simulated annealing algorithm implemented in the software Genodive 2.0b17 (Meirmans and Van Tienderen 2004) with 100000 steps over values of K ranging from 1-6 and examined the optimal value of K using the Calinski-Harabasz (1974) pseudo-F statistic and the Akaike Information Criterion (AIC). Calinski-Harabasz' pseudo-F statistic is thought to perform slightly better for population clustering (Meirmans 2010).

We also inferred the number of genetic clusters based on microsatellites. However, missing values in the microsatellite dataset could bias the clustering analyses if individuals missing data tend to be grouped together. To examine the effect of missing data on our analysis, we performed analyses on the original dataset containing missing values, as well as on two

additional datasets based on the original data, but where missing values were replaced with values drawn at random from the distribution of: 1) overall allele frequencies, and 2) population-specific allele frequencies. We used Genodive 2.0b17 (Meirmans and Van Tienderen 2004) to create these new datasets. We estimated *K* in Structure using a model that incorporates sample group information to assist clustering when population structure is weak (Hubisz et al. 2009). We explored values of *K* ranging from 1-12. Ten runs consisting of 100000 sweeps after a burn-in of 50000 sweeps were conducted for each value of *K*. We chose the best *K* following the method of Evanno et al. (2005). We also modeled *K* as a random variable in Structurama, using four chains of 100000 sweeps each sampled every 25 generations, with the first 1000 samples discarded as burn-in. We also used GenoDive with 100000 steps over values of *K* ranging from 1-12 and examined the optimal value of *K* using the pseudo-F statistic and AIC. We ran each analysis multiple times from different random starting points to ensure convergence. Structure Harvester v0.3 was used to summarize and view results from Structure output (Earl 2009). Distruct 1.1 (Rosenberg 2004) was used to plot individual membership coefficients from Structure.

ANALYSIS OF MOLECULAR VARIANCE

We used a hierarchical analysis of molecular variance (AMOVA) framework to investigate the relative contribution of variance among groups of populations in different refugia, among populations within refugia, and among individuals within populations (Excoffier et al. 1992). Populations not found in putative refugia were excluded from AMOVA analyses. We tested the null hypothesis of panmixia (one group) as well as alternative structuring by refugium (four groups). For the microsatellites, we used both the original dataset containing missing values as well as the two datasets in which missing values were replaced with values drawn from the distribution of overall allele frequencies or population allele frequencies. Genetic distances were calculated using an infinite allele model for the allozymes, stepwise mutation model for the microsatellites, and the Tamura and Nei (1993) method for the mtDNA. Significance of the covariance components was assessed using 20000 permutations, a sufficient number for having less than 1% difference with the exact probability in 99% of cases (Guo and Thompson 1992). AMOVAs for the allozymes and microsatellites were performed in GenoDive 2.0b17 (Meirmans and Van Tienderen 2004); the mtDNA AMOVA was performed in Arlequin 3.5.1.2 (Excoffier and Lischer 2010).

PHYLOGENETIC ANALYSIS

Phylogenetic analysis of mtDNA was conducted using Bayesian methods. The tree was rooted using *Ensatina eschscholtzii croceater*, the closest relative of *klauberi*. The mtDNA was used to conduct a Bayesian analysis partitioned by codon and tRNAs in MrBayes v3.1.2 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). The Akaike information criterion (AIC) implemented in MrModeltest v2 (Nylander 2004) was used to choose the best-fit nucleotide substitution model for each partition. GTR+Γ was used for first codon positions, HKY+I for second codon positions and tRNAs, and K80 for third codon positions. Four independent analyses were run for $2x10^7$ generations, each using random starting trees and default priors. In each analysis, four Markov chains (using default heating values) were sampled every 1,000 generations. Stationarity was evaluated by visually examining parameter trace files from MCMC runs in Tracer v1.5 (Rambaut and Drummond 2007). Trees sampled prior to reaching stationarity

were discarded as burn-in, and the remainder used to construct a 50% majority rule consensus tree. Posterior probability values were used to assess phylogenetic support (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003); clades with values greater than 95% were considered to be significantly supported.

TESTING THE EFFECTS OF CLIMATIC STABILITY ON POPULATION STRUCTURE AND CONNECTIVITY

To examine whether climatic stability has influenced population genetic structure, we first used a circuit-theoretic model to measure the connectivity among the centroids of areas identified as refugia using the stability surface as a conductance grid in the software Circuitscape v3.3 (McRae 2006; McRae and Beier 2007; McRae et al. 2008). We then tested whether historical population connectivity was correlated with gene flow by performing a partial Mantel test (Smouse et al. 1986) of pairwise Φ_{ST} and population connectivity while controlling for geographic distance. Significance of the correlation was assessed using 1000 permutations implemented in GenoDive 2.0b17 (Meirmans and Van Tienderen 2004).

Results

SPECIES DISTRIBUTION MODELING

The distribution model for current climatic conditions performed well over the majority of the species' range (area under the receiver operating characteristic curve (Hanley and McNeil 1982) value of 0.976). Two populations did not fall within the predicted range, the Sierra de Juarez and San Quintín samples (Fig. 2A). Populations at the northern and southern distributional limits of the species' range in the San Bernardino Mountains and Sierra San Pedro Mártir, respectively, showed a lower probability of occurrence relative to the San Jacinto Mountains, Palomar Mountain, and the Cuyamaca Mountains. There was some minor overprediction of the model into the San Gabriel Mountains west of the San Bernardino Mountains where the species is not known to occur. The HOL model (6 ka) predicted a broader, more continuous distribution than the current range (Fig. 2B). The LGM model (21 ka) predicted a more narrow distribution and southward shift in the northern range limit (Fig. 2C) of the species, excluding it from the San Bernardino Mountains. The stability map representing the intersection of the models across all three periods resulted in five separate refugia, the largest extending from Palomar Mountain southeast to the Cuyamaca and Laguna Mountains, and four smaller refugia, in the southern San Bernardino, San Jacinto, and Santa Rosa mountains, respectively (Fig. 2D).

DESCRIPTIVE STATISTICS

Descriptive statistics for the allozymes and microsatellites are provided in Appendixes 3 and 4, respectively. The allozymes showed relatively low levels of polymorphism. Significant departures from HWE were observed at two loci in the Palomar Mountain population and three loci in the Cuyamaca population. Microsatellite polymorphism was high. Significant departures from HWE were again found in the Palomar Mountain and Cuyamaca populations. Departures at

these microsatellite loci were due to null alleles, as in the Palomar population, or heterozygote deficit, as in the Cuyamaca population.

POPULATION GENETIC VARIATION AND STRUCTURE

Pairwise comparisons of F_{ST} were significant in most comparisons for the allozymes (Table 1) and microsatellites (Table 2). Pairwise differences for the mtDNA are provided in Table 3. The highest levels of genetic differentiation based on allozymes were between northern (Crystal Creek, Sawmill Canyon, Fuller Mill Creek) and southern (Palomar and Cuyamaca mountains) populations (Table 1). Levels of differentiation were lower for the microsatellites, but the San Quintín population showed high levels of differentiation from all other populations (Table 2). Clustering analyses were concordant with these results. When the number of clusters based on allozymes was considered to be fixed over a range of values from 1-6 in Structure, the highest posterior probability was obtained for $K = 2$ using the mean log-likelihood (Pritchard et al. 2000) or ΔK (Evanno et al. 2005) as criteria (Fig. 4). When the number of populations was modeled as a random variable following a Dirichlet process prior in Structurama, the mean population partition was $K = 2$ (average squared distance to partition = 2.0565). Both methods resulted in identical individual assignments to the two groups. Populations are structured geographically into a northern group consisting of the San Bernardino (Crystal Creek, Sawmill Canyon), San Jacinto (Fuller Mill Creek), and Santa Rosa mountains (Queen Creek), and a southern group consisting of the Palomar and Cuyamaca Mountain populations (Fig. 5).

Clustering analysis of the microsatellite data resulted in two or 5-7 groups, depending on the criterion used to choose the optimal number of groups and whether missing data were included in the analysis (Table 4). The San Quintín population was always recovered as a distinct cluster. Using the pseudo-F statistic, the optimal number of clusters was $K = 2$, in general agreement with results from Struture (Table 4). Using the AIC resulted in 5-7 clusters, depending on whether the original dataset with missing values was used ($K = 6$), the dataset filled using random alleles drawn from the overall allele frequency distribution ($K = 5$), or the dataset filled using alleles drawn from the population allele frequency distribution ($K = 7$) (Table 4).

ANALYSIS OF MOLECULAR VARIANCE

The hierarchical AMOVA revealed significant genetic variation for all three marker classes, allowing us to reject the null hypothesis of panmixia (Table 5). Most of the total nuclear genetic variance was found within populations, while most of the total mitochondrial variance was found among populations (Fig. 8). Geographic structuring of variation by refugia was not supported by the AMOVA for any of the three datasets.

PHYLOGENETIC DATA ANALYSIS

There were 29 unique mtDNA haplotypes among the 70 individuals sequenced. Average maximum-likelihood corrected pairwise sequence divergence among ingroup haplotypes was 1.3%, with a maximum of 5.4%. Average sequence divergence between *E. e. klauberi* and the outgroup *E. e. croceater* was 8.3%. The deepest division separated a Sierra San Pedro Mártir clade from the remainder of the sample (Fig. 9). Within the larger clade, the next deepest division separated two individuals from the north side of the San Bernardino Mountains from the

rest of the tree, although a third haplotype from nearby Crystal Creek in the San Bernardinos grouped with individuals from the San Jacinto Mountains to the south. Individuals from the San Quintín population formed a well-supported clade, sharing a single haplotype. The single individual from the Sierra de Juarez in northern Baja California grouped with individuals to the north in the San Diego ranges rather than the Sierra San Pedro Mártir to the south. There is some further local geographic structure, but relationships among the tips of the tree are not well-supported.

EFFECTS OF CLIMATIC STABILITY ON POPULATION STRUCTURE AND CONNECTIVITY

Partial Mantel tests of historical population connectivity and pairwise population divergence (Φ_{ST}) controlling for geographic distance were not significant for allozymes ($p = 0.615$), microsatellites ($p = 0.588$) or mitochondrial DNA ($p = 0.247$).

Discussion

Here, we have integrated an ecological niche model constructed from climatic characteristics of known species occurrence with population genetic and phylogeographic information to understand the effects of late Quaternary climatic oscillations on the distribution and genetic structure of a dispersal-limited terrestrial salamander. Because terrestrial plethodontids are so intimately linked to the environment physiologically (Feder 1983), they make excellent study organisms for such an approach. We have documented "refugial" areas that hypothetically may have harbored mesic-adapted species during drier periods, providing spatially explicit demographic hypotheses that we tested with multilocus genetic data. This approach has proven to be particularly powerful in phylogeography (Carstens and Richards 2007; Richards et al. 2007; Waltari et al. 2007; Carnaval et al. 2009), but is not without its limitations (Peterson and Nyári 2008). One of the greatest difficulties comes from issues with spatial resolution. We are using global climate models that may be too coarse for the relatively fine-scale resolution of our study. Additionally, when distribution models developed for present climatic conditions are projected back in time to time periods where conditions existed that do not currently, results may be misleading (Pearson et al. 2006). Finally, these "refugia" are hypothetical only; when dispersal is limited, some refugia may have been inaccessible to a species (Araújo and Pearson 2005). These limitations must be considered when interpreting results.

POPULATION GENETIC VARIATION AND STRUCTURE

Limitations of F-statistics notwithstanding (Neigel 2002), pairwise population comparisons of genetic differentiation support significant population subdivision. The hierarchical AMOVAs showed significant genetic structuring, however most of the total genetic variance was found within, rather than among, populations. There is some discordance between nuclear and mitochondrial evidence, with mitochondrial data showing higher genetic variation among geographic regions compared to the allozymes and microsatellites.

Significant departures from HWE were observed in the Palomar Mountain and Cuyamaca populations for both the allozymes and microsatellites. Departures from HWE in the microsatellites were due to the presence of null alleles, as in the Palomar population, and/or

heterozygote deficit, as in the Cuyamaca population. Heterozygote deficit in these populations may also be attributable to subpopulation structure (the Wahlund effect) given that neighboring samples were pooled from several adjacent localities for both the Palomar and Cuyamaca populations.

Consistent with pairwise population comparisons, results from Structure and Structurama based on allozymes inferred two genetic clusters, a northern group made up of populations in the San Bernardino, San Jacinto, and Santa Rosa mountains, and a southern group, made up of populations in the Palomar, Volcan, Cuyamaca, and Laguna mountains of San Diego County (hereafter, the "San Diego ranges"). These two groups are separated by a valley coinciding with the northwest-trending San Jacinto Fault along the south side of the San Jacinto (and Santa Rosa) Mountains (Hall 2007). Clustering based on the microsatellite data did not recover this break, but revealed that the San Quinítin population was genetically distinct to the exclusion of the remainder of the sample.

EFFECTS OF LATE QUATERNARY CLIMATE CHANGE ON POPULATION STRUCTURE

Geographic structuring of genetic variation by putative refugia was not supported by the AMOVA, suggesting that populations in these "stable" areas have not been isolated for very long, or that recurrent patterns of gene flow in response to cyclical environmental changes and habitat shifts may have masked any earlier periods of divergence in allopatry. Below, we briefly summarize major features of climate change in California to provide background for understanding patterns of population connectivity and genetic structure in the Large-blotched Ensatina and other mesic-adapted species inhabiting montane regions of southern California.

California's modern mediterranean climate of winter precipitation and summer drought (formerly tropical) was established during the Tertiary as a result of seasonal changes in global circulation from "greenhouse" to "icehouse" states (Minnich 2007). Summer drought developed by the Miocene (Minnich 2007). Since the beginning of the Pleistocene around 2.6 Ma, California's climate has fluctuated with glacial-interglacial cycles. During glacial periods, levels of precipitation were greater than at present; during interglacials, climate was similar to the present one (Minnich 2007). The last major glacial period, the Wisconsin glaciation, began about 121 ka. There have been about six minor pluvial-interpluvial cycles within that period, the most recent one being the most extreme (Schaffer 1993). There remains disagreement as to what conditions in California were like under a full-glacial climate (Minnich 2007). Of the very few macrofossil floras that exist, the early Pleistocene (1 Ma) Soboba flora from near San Jacinto (Riverside County) suggests that minimum precipitation was 63.5 cm (ca. 25 in.) and the area contained riparian woodland, bigcone Douglas fir forest, mixed conifer forest, and chaparral (Axelrod 1966). Today, the area that contained the fossil flora is semi-desert (~28 cm or 11 in. of precipitation) (Axelrod 1989).

Evidence of late Quaternary vegetation changes in California comes primarily after about 25 ka (Minnich 2007), much of which is based on studies of pollen and packrat middens (Spaulding 1990; Van Devender 1990). The last glacial maximum (LGM) took place from approximately 21-16 ka (Bartlein et al. 1998). The climate in California was cooler and wetter, as evidenced by glaciation in the Sierra Nevada and pluvial lakes in lowland areas of the Great Basin (Minnich 2007). The southwestern-most limit of glaciation in the Western Cordillera occurred in the San Bernardino Mountains, with glaciers reaching their greatest extent there 20-16 ka (Owen et al. 2003).

The transition from the Pleistocene to the modern Holocene climate 11-10 ka marked a shift to a warmer climate in California along with deglaciation (Minnich 2007). Temperatures reached their maximum by 6800 years before present (BP) (Feng and Epstein 1994), but glaciers were still present in the San Bernardino Mountains as recently as the early-middle Holocene (5-9 ka) (Owen et al. 2003). The mid-Holocene in California (5-3.5 ka) was a period of greater precipitation, expansion of Sierra Nevada glaciers, and formation of small pluvial lakes at lower elevations, due to increased frequency of moist westerlies along the Pacific Coast (Minnich 2007). Several areas of California show a shift to a more mesic vegetation (Minnich 2007). The late Holocene in California in contrast was characterized by a period of reduced precipitation from 2000-600 BP and major drought beginning about 1000 BP (Minnich 2007). There is little, if any, evidence for the effects of this drought on forests, though presumably mixed conifer forest moved upslope. Desert vegetation appears to have moved upslope by 100-200 m between 1500-500 BP (Spaulding 1990).

At first glance, the relatively low levels of genetic differentiation between populations of Large-blotched Ensatina are somewhat surprising, given the limited dispersal ability of *Ensatina* (females: max. 60.6 m, mean 23.3; males: max. 120.4 m, mean 31.2 m (Staub et al. 1995)) and high levels of differentiation found within other subspecies of the *Ensatina* complex (Wake and Yanev 1986; Kuchta et al. 2009a; Pereira and Wake 2009). Our species distribution model for the LGM predicts a substantial reduction in range compared to the present, failing to predict species presence in the formerly glaciated San Bernardino Mountains (Fig. 2C). Our HOL model predicts a much larger distribution and greater connectivity among populations (Fig. 2B), suggesting that increased precipitation and temperature at this time may have allowed for dispersal into low elevation areas increased gene flow among populations. Some populations have persisted at low elevations when there is sufficient moisture available, such as in areas near springs on the north slope of the San Bernardino Mountains in semi-desert habitat, or even in coastal areas at sea-level that receive winter rain and summer fog, such as the population near San Quintín. The physiognomy of the area inhabited by this localized population stands in stark contrast to the high-elevation conifer and mixed conifer forest inhabited by all other known populations; habitat at the collecting locality consists of a treeless rocky area surrounded by maritime desert scrub.

CONSERVATION IMPLICATIONS

Like many other plants and animals endemic to the California Floristic Province (a biodiversity hotspot (Myers 2000)), this species is increasingly threatened by development, fires, agriculture, logging, and climate change (Stephenson and Calcarone 1999; Minnich and Franco-Vizcaíno 2005). Evidence from plant distributions in Central California show that major shifts in vegetation may result from as little as 1° C change in mean temperature of the warmest or coldest month (Axelrod 1981). Distribution modeling under future (2x current CO_2 levels) climatic conditions predicts a significant contraction in this species' range (not shown). Owing to its limited, fragmented distribution in densely populated southern California, the Large-blotched Ensatina is classified as a California Species of Special Concern and a Forest Service Region 5 Sensitive Species (Stephenson and Calcarone 1999) due to significant habitat loss and conversion. The mixed conifer forest of the Peninsular Range "sky-islands" inhabited by this species in Baja California represent the only multi-species, mediterranean-climate forests in all of Mexico (Ferrusquía-Villafranca 1993) and is under increased threat from fires, cattle ranching

and logging (Minnich and Franco-Vizcaíno 2005). The coastal population in Baja California is particularly vulnerable to habitat loss given its proximity to the city of San Quintín, an area that is rapidly being developed.

Acknowledgements

We thank Craig Moritz, Jimmy A. McGuire, David B. Wake, and George K. Roderick for comments on an earlier draft of this manuscript. We thank Monica Frelow and Kay Yanev for access to their unpublished allozyme data. For collecting and donating samples, we thank Chuck Brown, Dave Goodward, Ted Papenfuss, Jim Parham, Brad Hollingsworth, Jorge Valdez-Villavicencio, and Anny Peralta-Garcia. For assistance in the field, we thank Chuck Brown, Greg Pauly, Megan Lahti, Jorge Valdez-Villavicencio, Clark Mahrdt, Jim McGuire, Brad Hollingsworth, and Angelo Soto-Centeno. For arranging permits and logistics to work in Mexico, we thank Brad Hollingsworth. Funding was provided by a Howie Wier Conservation Award from the Anza-Borrego Institute, a University of California Natural Reserve System (UCNRS) Mildred E. Mathias graduate student research grant, the Museum of Vertebrate Zoology Martens and Louise Kellogg funds, the San Diego Natural History Museum (SDNHM), the University of California Institute for Mexico and the United States (UC MEXUS), and the National Science Foundation (Doctoral Dissertation Improvement Grant DEB-0909821 and NSF DEB-0641078). We thank the ESRI Conservation Program for providing ArcGIS software for our use.

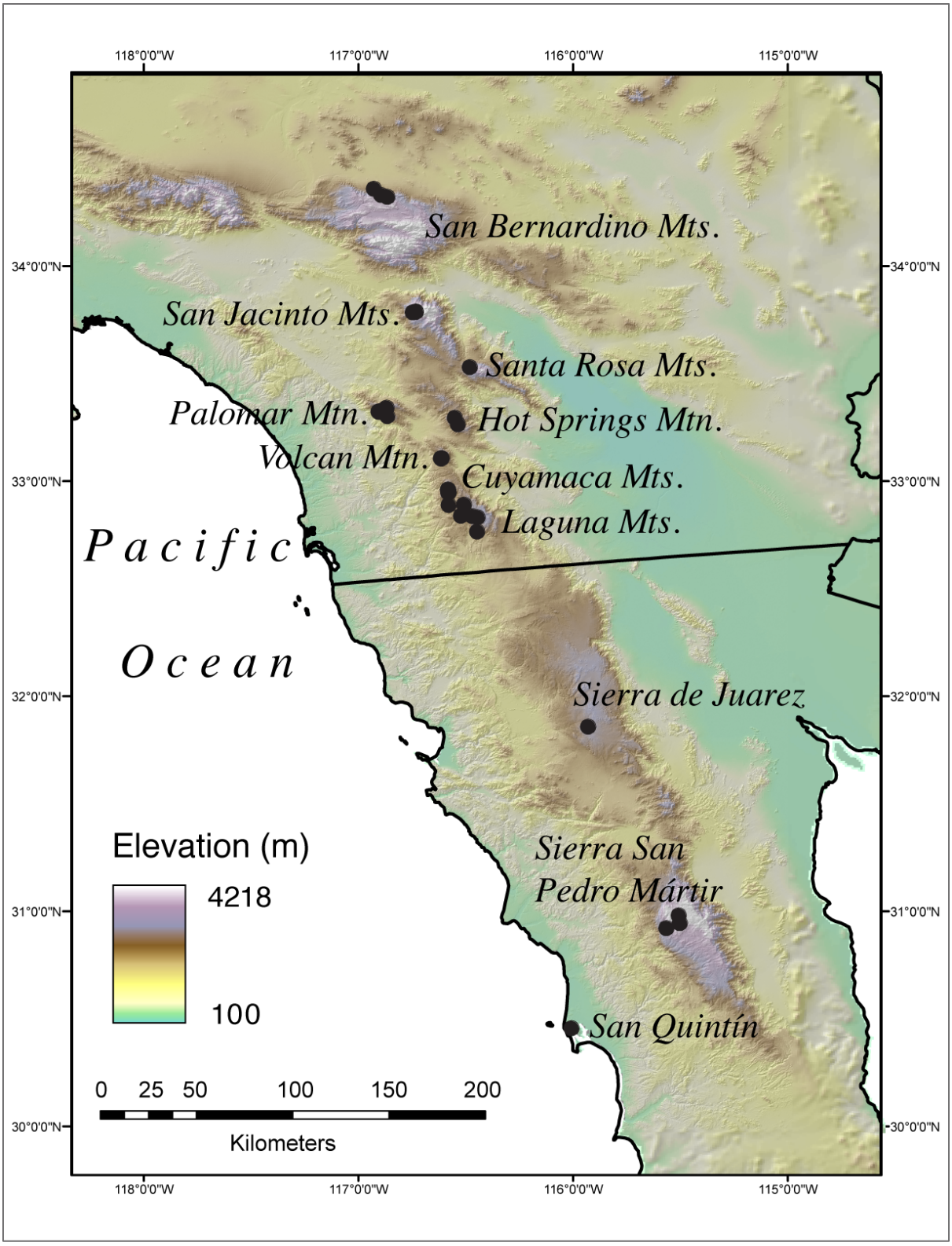

Figure 1. Distribution of the Large-blotched Ensatina (*Ensatina eschscholtzii klauberi*) in southern California and northern Baja California showing localities used in distribution modeling.

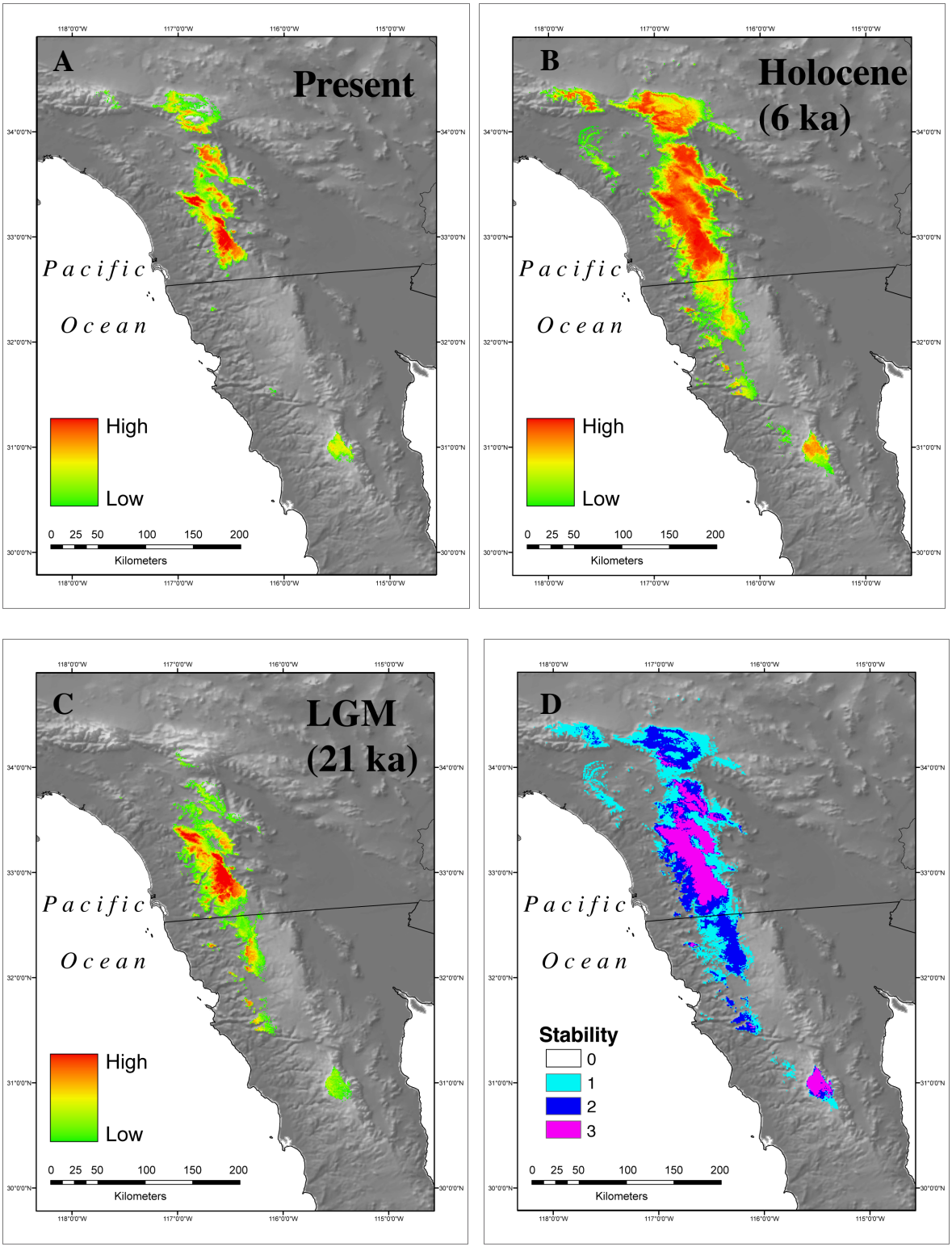

Figure 2. Species distribution models at present (A), Holocene (HOL, 6 ka; B), Last Glacial Maximum (LGM, 21 ka; C), and a stability map representing the intersection of the models over all three time periods (D). Stability values range from 0 (never present) to 3 (present during all three time periods).

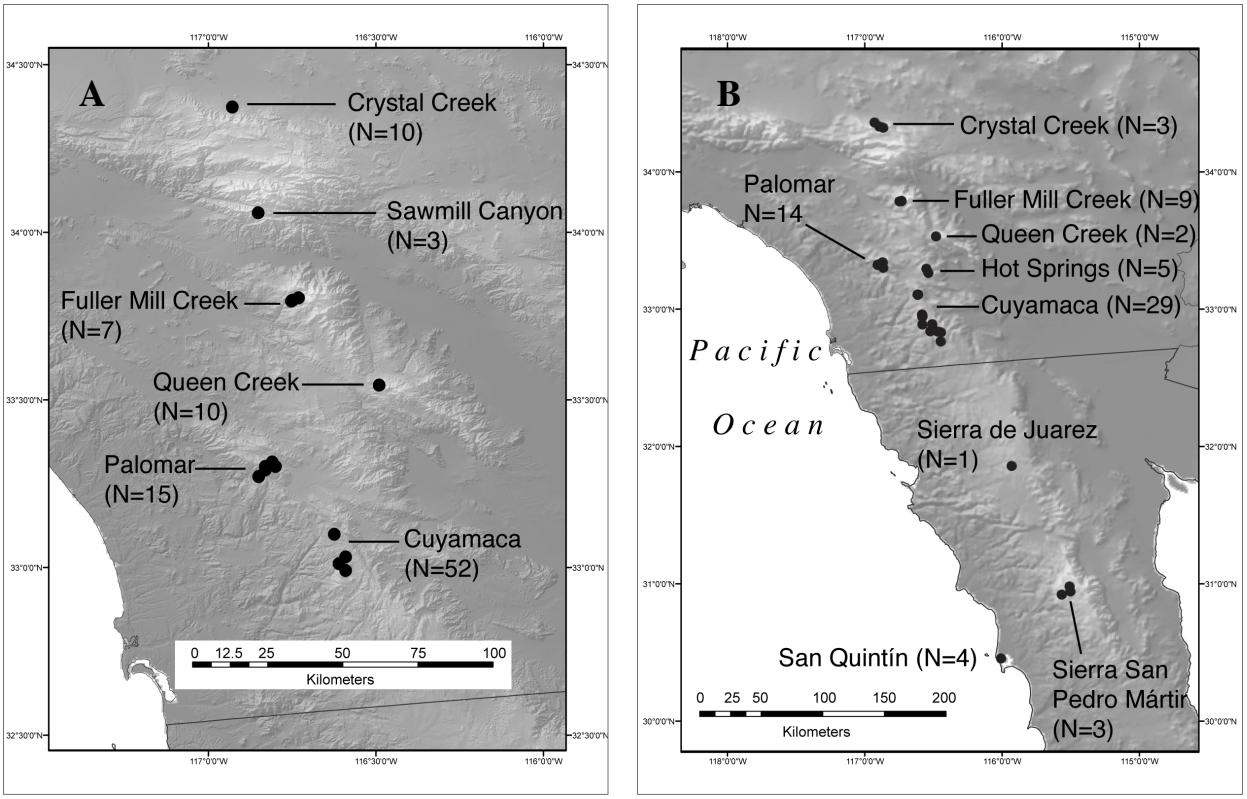

Figure 3. Sampling for allozymes (A), and microsatellites and mitochondrial DNA (B).

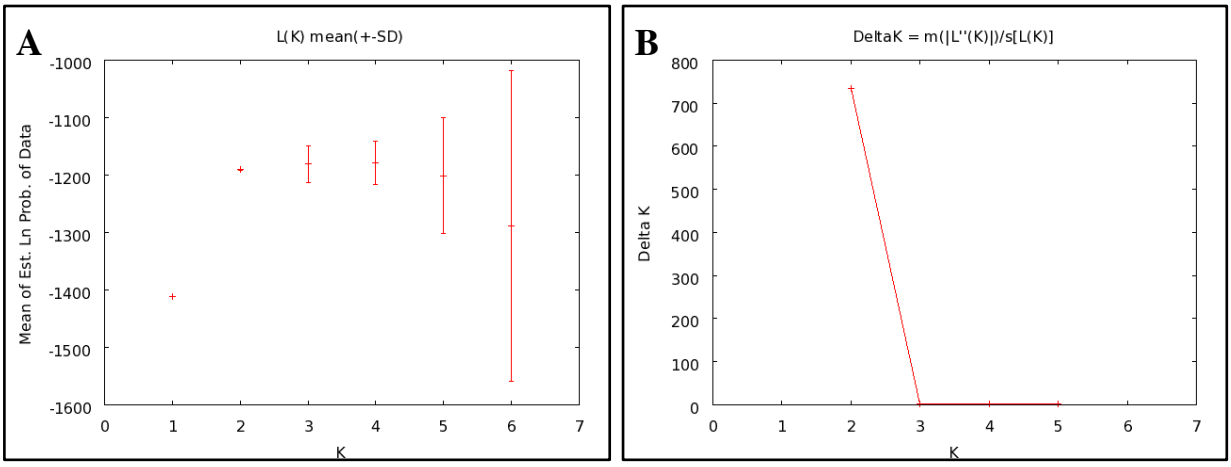

Figure 4. Number of Large-blotched Ensatina populations with the highest posterior probability based on 19 allozyme loci over 10 Structure runs for each value of *K* from one to six expressed as: A) the mean log likelihood ($L(K) \pm SD$); B) ΔK calculated as $\Delta K = m\left|L''(K)\right|/s[L(K)]$. The mode of this distribution is considered the best estimate of the level of genetic structure (Evanno et al. 2005), here two clusters. Plots were created using Structure Harvester v0.3 (Earl 2009).

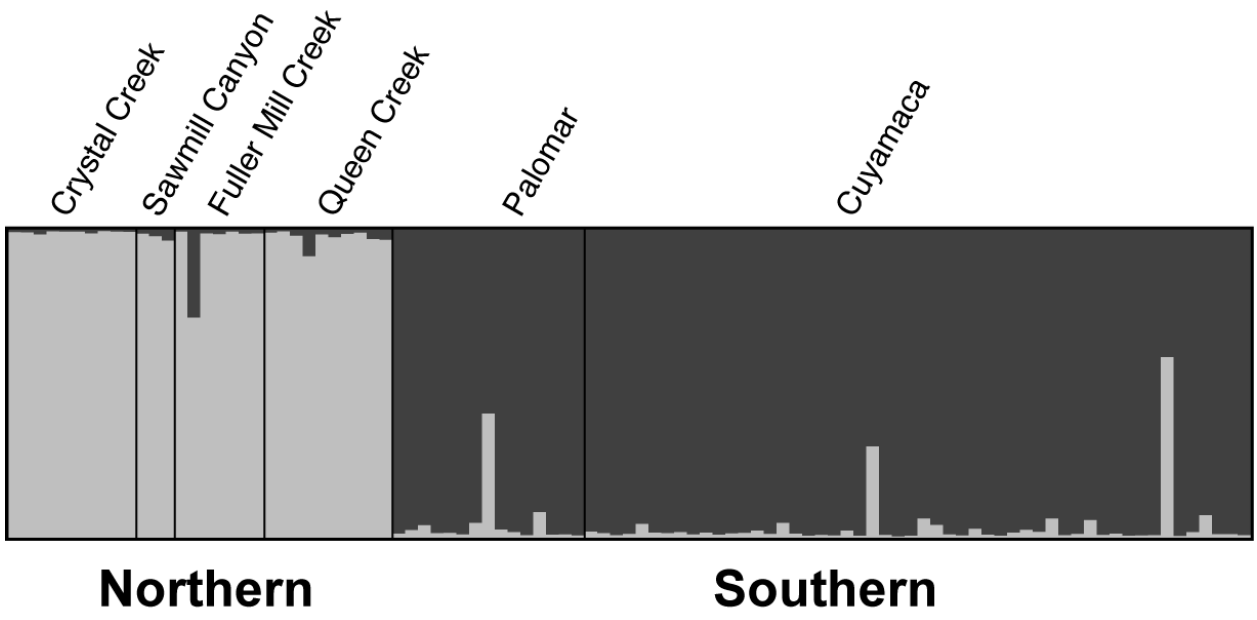

Figure 5. Number of Large-blotched Ensatina populations (*K* = 2) based on 19 allozymes.

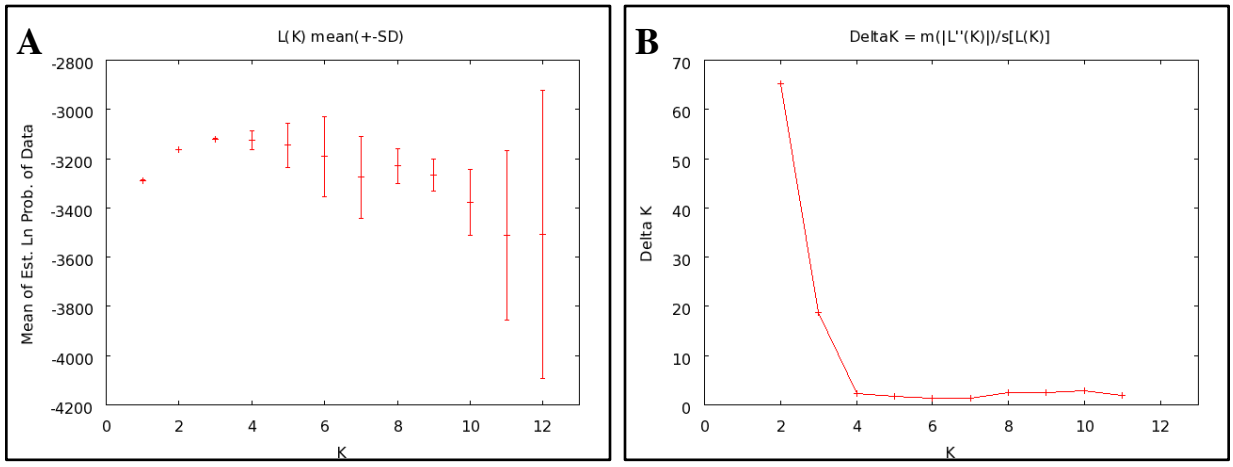

Figure 6. Number of Large-blotched Ensatina populations with the highest posterior probability based on 10 microsatellite loci over 10 Structure runs for each value of *K* from 1-12 expressed here as the mean log likelihood ($L(K) \pm$ SD) on the left and ΔK calculated as $\Delta K = m\left|L''(K)\right|/s[L(K)]$ on the right (Evanno et al. 2005). The mode of the ΔK distribution is considered the best estimate of the level of genetic structure, here two clusters. Plots were created using Structure Harvester v0.3 (Earl 2009).

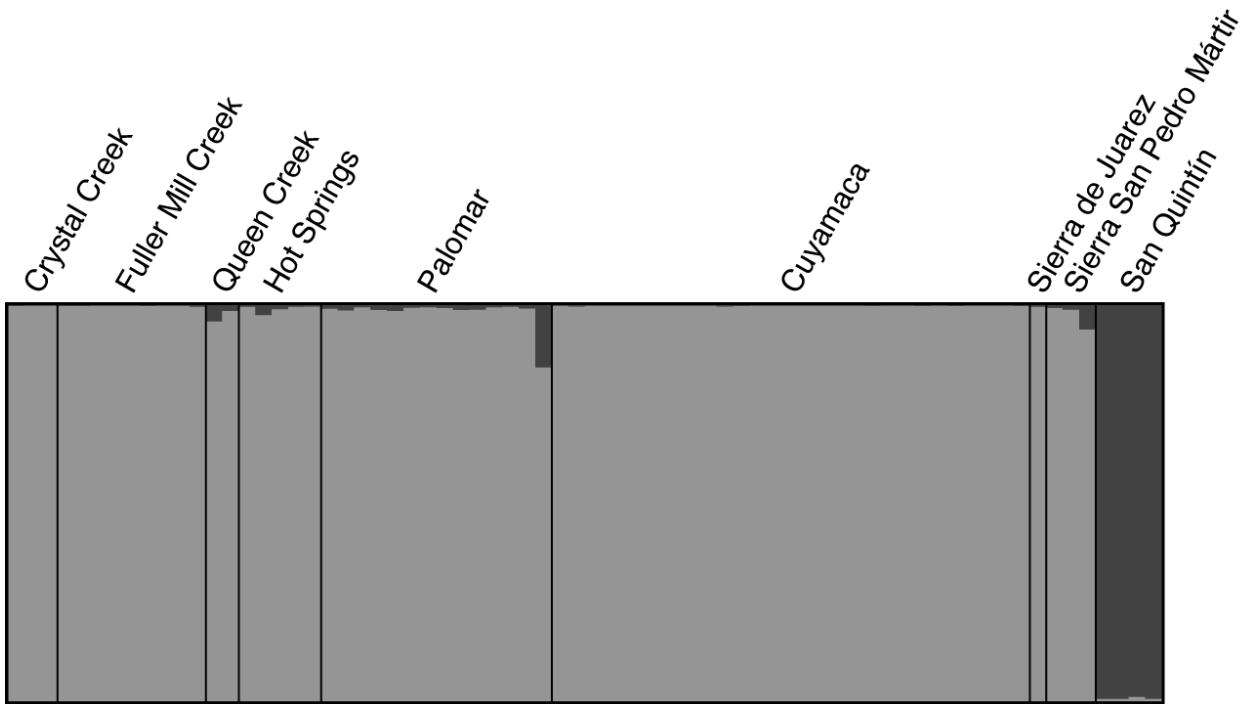

Figure 7. Number of Large-blotched Ensatina populations based on the original microsatellite dataset for *K* = 2.

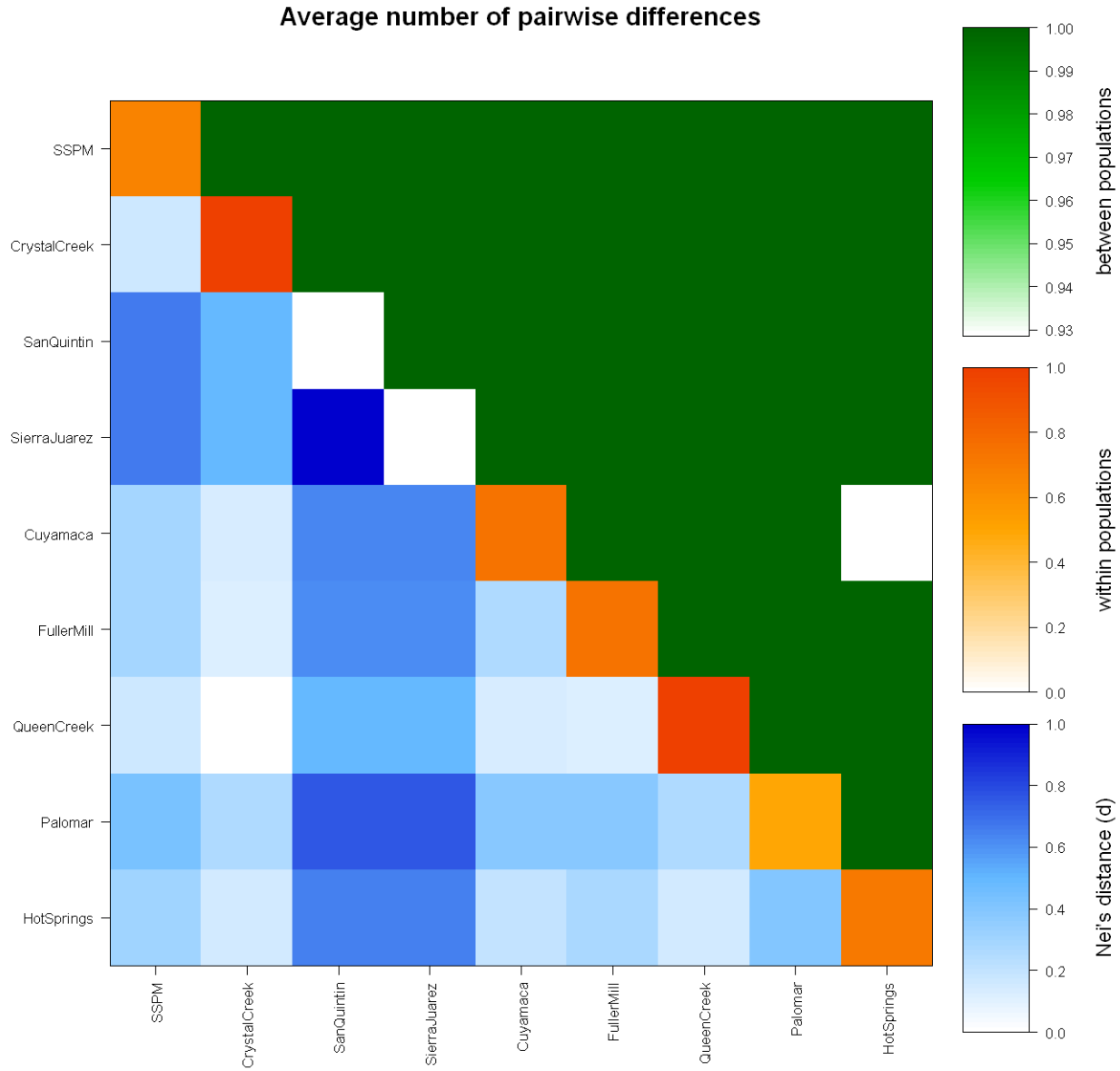

Fig. 8. Average number of pairwise differences (π) within and between populations based on mtDNA. Orange colors on the diagonal represents π within populations; green colors above the diagonal represent π_{XY} between pairs of populations; blue colors below the diagonal represent the net number of nucleotide differences between populations (D_A).

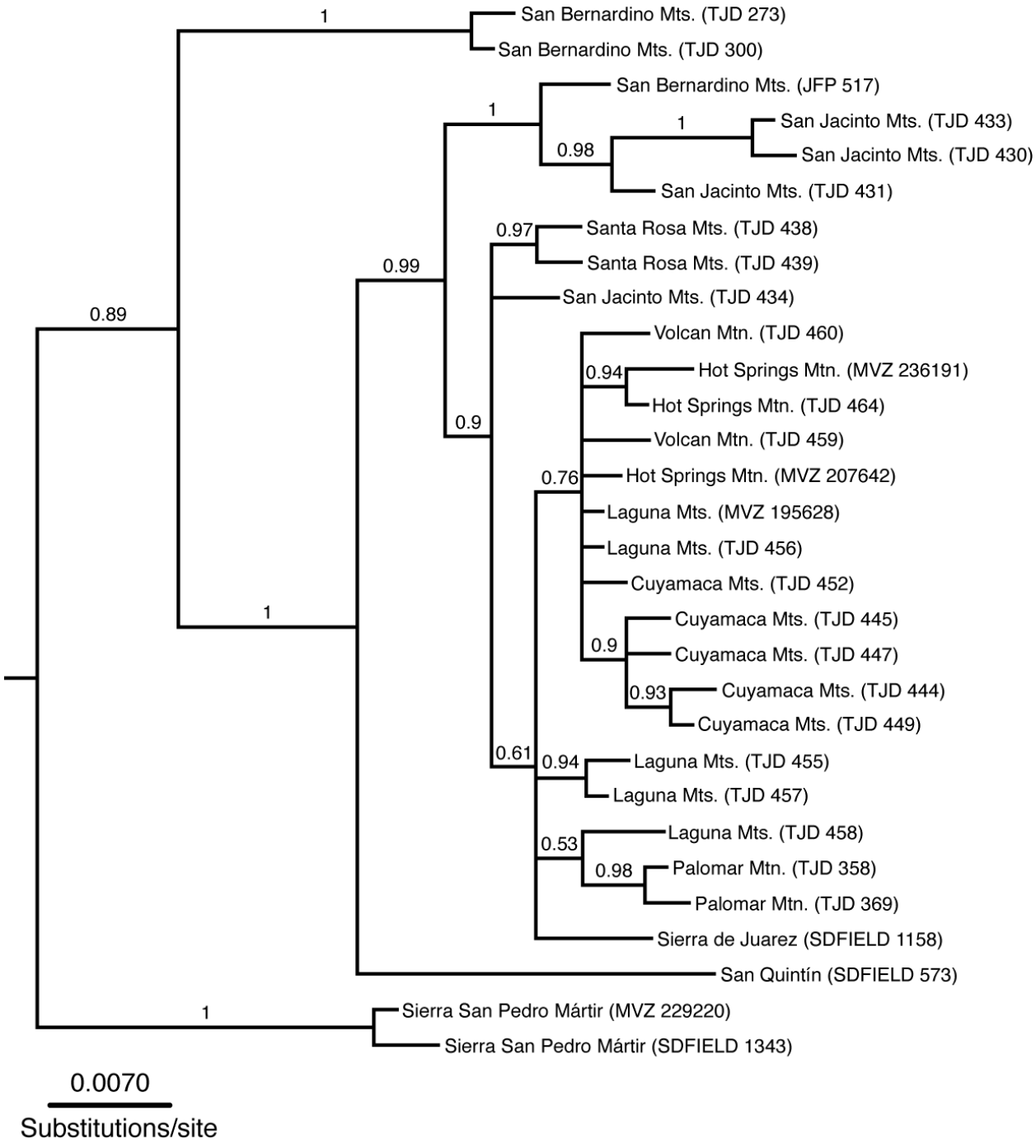

Figure 9. Phylogenetic relationships of based on the 50% majority rule consensus tree from Bayesian analysis of unique haplotypes only. Numbers on branches represent Bayesian posterior probability values. Field numbers correspond to those in Appendix 2.

Table 1. Genetic distances (Nei's *D*, below diagonal; pairwise F_{ST} above diagonal) between six populations of the Large-blotched Ensatina from 19 allozyme loci. Numbers in bold indicate significant F_{ST} p-values based on 20000 permutations.

	Crystal	Sawmill	Fuller Mill	Queen	Palomar	Cuyamaca
Crystal	—	**0.25**	**0.31**	**0.39**	**0.50**	**0.43**
Sawmill	0.03	—	0.14	**0.22**	**0.38**	**0.29**
Fuller Mill	0.05	0.03	—	0.03	**0.31**	**0.28**
Queen	0.08	0.05	0.01	—	**0.30**	**0.26**
Palomar	0.12	0.09	0.07	0.07	—	**0.07**
Cuyamaca	0.14	0.08	0.08	0.07	0.01	—

Table 2. Genetic distances (R_{ST} below diagonal, F_{ST} above diagonal) between nine populations of the Large-blotched Ensatina based on the three microsatellite loci with the most complete data (ENS1, ENS6, ENS7). Numbers in bold indicate significant F_{ST} p-values based on 20000 permutations.

	Crystal	Fuller Mill	Queen	Palomar	Hot Spgs.	Cuyamaca	S. Juarez	SSPM	San Quintín
Crystal	—	**0.07**	0.27	**0.15**	**0.21**	**0.13**	0.30	0.20	**0.52**
Fuller Mill	0.08	—	**0.13**	**0.05**	**0.03**	0.01	0.08	**0.07**	**0.31**
Queen	0.38	0.15	—	**0.15**	0.08	0.09	0.30	0.21	0.56
Palomar	0.18	0.05	0.18	—	**0.06**	**0.06**	0.13	**0.15**	**0.27**
Hot Springs	0.27	0.04	0.09	0.07	—	0.01	0.12	**0.15**	**0.40**
Cuyamaca	0.15	0.01	0.10	0.06	0.02	—	0.07	**0.10**	**0.29**
S. Juarez	0.43	0.09	0.42	0.14	0.13	0.08	—	0.13	0.63
SSPM	0.24	0.08	0.27	0.18	0.17	0.11	0.14	—	**0.40**
San Quintín	1.10	0.45	1.26	0.38	0.66	0.42	1.69	0.65	—

Table 3. Pairwise distances between nine populations of the Large-blotched Ensatina based on 820 bp mtDNA. Numbers above diagonal are the average number of pairwise differences between populations, diagonal numbers are the average number of pairwise differences within populations, and numbers below diagonal are Tamura and Nei (1993) corrected average number of pairwise differences between populations.

	Crystal	Fuller Mill	Queen	Palomar	Hot Springs	Cuyamaca	S. Juarez	SSPM	San Quintín
Crystal	17.8	21.6	18.1	21.3	29.1	19.7	27.6	18.8	19.3
Fuller Mill	10.0	5.4	11.2	14.3	33.1	12.0	23.7	12.5	12.0
Queen	8.2	7.5	2.0	7.1	28.7	4.9	19.5	5.4	4.9
Palomar	12.3	11.6	6.1	0.0	27.6	6.2	20.5	6.4	5.6
Hot Springs	19.5	29.7	27.1	27.0	1.3	27.5	29.9	28.0	27.1
Cuyamaca	10.1	8.7	3.2	5.5	26.1	1.3	19.7	4.5	2.3
S. Juarez	18.7	21.0	18.5	20.5	29.2	19.1	0.0	20.2	19.9
SSPM	9.6	9.6	4.1	6.1	27.1	3.6	19.9	0.5	4.2
San Quintín	9.4	8.2	2.9	4.6	25.4	0.6	18.8	2.9	2.1

Table 4. The number of genetic clusters inferred based on allozymes and microsatellites. Sampling differs between the two marker types. Values for three versions of the microsatellite dataset are presented: the original microsatellite dataset with missing values, the dataset where missing values have been filled with alleles drawn at random from the overall allele frequency distribution, and the dataset where missing values have been filled with alleles drawn from the population-specific allele frequency distribution.

Dataset	STRUCTURE	STRUCTURAMA	Pseudo-F	AIC
Allozymes	$K = 2$	$K = 2$	$K = 2$	$K = 2$
Msats original (with missing values)	$K = 2$	$K = 6$	$K = 2$	$K = 6$
Msats filled-in with overall alleles	$K = 3$	$K = 4$	$K = 2$	$K = 5$
Msats filled-in with population alleles	$K = 2$	$K = 4$	$K = 2$	$K = 7$

Table 5. Results from hierarchical analysis of molecular variance (AMOVA) of genetic variation. Significance of the variance components and associated Φ-statistics was assessed using 20000 permutations. Genetic distances were calculated using an infinite allele model for the allozymes, a stepwise mutation model for the microsatellites, and using the Tamura and Nei (1993) method for the mtDNA. Standard errors of Φ-statistics for the nuclear data were obtained through jackknifing over loci. P-values shown in bold are significant.

Source of Variation	% Variance	Fixation Index	Standard error	*p*
Allozymes				
1. One group (Sawmill, Fuller Mill, Queen, Palomar, Cuyamaca)				
Within individual	0.718	$\Phi_{IT} = 0.282$	0.087	—
Among individual	0.072	$\Phi_{IS} = 0.091$	0.050	**0.005**
Among population	0.210	$\Phi_{ST} = 0.210$	0.066	**0.000**
2. Four groups (Sawmill) vs. (Fuller Mill) vs. (Queen) vs. (Palomar, Cuyamaca)				
Within individual	0.666	$\Phi_{IT} = 0.334$	0.104	—
Among individual	0.067	$\Phi_{IS} = 0.091$	0.050	**0.006**
Among population	0.057	$\Phi_{SC} = 0.072$	0.025	**0.000**
Among groups	0.210	$\Phi_{CT} = 0.210$	0.098	0.615
Microsatellites				
1. One group (Fuller Mill, Queen, Palomar, Hot Springs, Cuyamaca, SSPM)				
Within individual	0.819	$\Phi_{IT} = 0.181$	0.057	—
Among individual	0.032	$\Phi_{IS} = 0.038$	0.076	0.244
Among population	0.149	$\Phi_{ST} = 0.149$	0.040	**0.000**
2. Four groups (Fuller Mill) vs. (Queen) vs. (Palomar, Hot Springs, Cuyamaca) vs. (SSPM)				
Within individual	0.804	$\Phi_{IT} = 0.196$	0.055	—
Among individual	0.032	$\Phi_{IS} = 0.038$	0.076	0.239
Among population	0.125	$\Phi_{SC} = 0.130$	0.053	**0.000**
Among group	0.040	$\Phi_{CT} = 0.040$	0.103	0.699
mtDNA				
1. One group (Fuller Mill, Queen, Palomar, Hot Springs, Cuyamaca, SSPM)				
Among individual	0.219	—	—	
Among population	0.781	$\Phi_{ST} = 0.781$	—	**0.000**
2. Four groups (Fuller Mill) vs. (Queen) vs. (Palomar, Hot Springs, Cuyamaca) vs. (SSPM)				
Among individual	0.144	$\Phi_{SC} = 0.502$	—	**0.000**
Among population	0.145	$\Phi_{ST} = 0.856$	—	**0.000**
Among group	0.711	$\Phi_{CT} = 0.711$	—	0.051

Chapter 3

Multilocus Phylogeography of the Salamander Ring Species *Ensatina eschscholtzii*

Abstract

Salamanders of the *Ensatina eschscholtzii* complex have featured prominently in evolutionary biology because they represent one of the few examples of a ring species. Decades of genetic work based on allozymes and mitochondrial DNA (mtDNA) have revealed remarkable levels of geographically structured genetic diversity and generally support the original ring species biogeographic history first described by Stebbins. Thus far, however, phylogenetic relationships have been inferred using only mtDNA, which may be problematic if the mitochondrial genealogy (the gene tree) differs from the true relationships between lineages (the species tree). Here, we use multiple nuclear DNA sequence based markers and one mitochondrial gene using both concatenation and species tree methods to reexamine patterns of lineage divergence underlying the historical biogeographic ring species scenario. Our results are generally consistent with previous efforts based on mtDNA, but build on these works by providing further resolution to the *Ensatina* phylogeny, while highlighting the difficult nature of inferring species trees from samples of closely related populations that are experiencing gene flow.

Introduction

The fields of phylogeography and phylogenetics have entered an exciting new era in which multilocus approaches are increasingly being used to reconstruct phylogenies (Brito and Edwards 2009; Edwards 2009). This surge in the availability of large molecular datasets poses new and interesting analytical challenges. One particularly salient difficulty is reconciling different genealogical histories (gene trees) to infer the true tree of relationships between species (the species tree). It has long been recognized that individual gene trees are expected to differ from one another and will not necessarily reflect the underlying species tree (Tajima 1983; Pamilo and Nei 1988; Maddison 1997; Nichols 2001). Gene trees differ from species trees for a variety of reasons, including horizontal transfer, incomplete lineage sorting (deep coalescence), and gene duplication and extinction (reviewed in Maddison 1997).

Diverse new methods for inferring the species tree when there is discord among gene trees have recently become available (reviewed in Degnan and Rosenberg 2009). The choice of the best method for species tree inference is largely data set dependent, and will vary based on population history, number of taxa, and number of characters (Knowles 2009; McCormack et al. 2009). Coalescent theory (Kingman 1982, 2000; Wakeley 2008) provides an intuitive conceptual framework for species tree inference. The coalescent follows individual gene copies backward in time from the present to the point that they merge with their most recent common ancestor. Extended from coalescent theory, the 'multispecies coalescent' (Degnan and Rosenberg 2009) is a probabilistic model that generalizes the Wright-Fisher model of genetic drift by applying it to multiple populations connected by an evolutionary tree. Gene trees can be easily simulated under

this model, which provides a useful guide for examining causes of gene tree discordance (Degnan and Rosenberg 2009).

The *Ensatina eschscholtzii* complex of plethodontid salamanders has received special attention from evolutionary biologists because it represents one of the few examples of a ring species (Futuyma 1998; Irwin et al. 2001). A ring species forms when a single species expands its range around a central geographic barrier along two separate pathways, with the two terminal forms gradually diverging and eventually becoming reproductively isolated where they meet and reconnect on the other side of the barrier. In his detailed analysis of geographic variation and speciation in *Ensatina* Stebbins (1949) concluded that the four previously recognized species of *Ensatina* (*eschscholtzii*, *sierrae*, *croceater*, and *platensis*) actually represented a single polytypic species comprising seven well-defined subspecies based primarily on differences in color pattern. Stebbins hypothesized that the complex originated in northwestern North America, probably in northern California and southern Oregon, where an ancestral population divided into two subpopulations that expanded southward around the arid Central Valley of California along two separate paths, one along the relatively low-elevation coast ranges, the other inland along the western slopes of the higher-elevation Sierra Nevada. Eventually, the two lineages came back into contact in southern California, creating a ring encircling the Central Valley.

Decades of analysis of both allozymes and mitochondrial DNA (mtDNA) have provided a fine-scale view of spatial patterns of genetic differentiation in *Ensatina*, revealing high genetic diversity and a more complex biogeographic history characterized by periods of vicariant allopatric divergence followed by secondary contact (Wake and Yanev 1986; Moritz et al. 1992; Kuchta et al. 2009b; Pereira and Wake 2009). Different interpretations of these results have generated controversy regarding species delimitation and the most appropriate taxonomy (Frost and Hillis 1990; Graybeal 1995; Wake and Schneider 1998) with some workers proposing that 11 or more species should be recognized (Highton 1998).

To date, phylogeographic relationships within the *Ensatina* complex have been based on mitochondrial DNA (mtDNA) gene genealogies (gene trees) (Moritz et al. 1992; Kuchta et al. 2009b). Moritz et al. (1992) provided the first mtDNA phylogeny of the *Ensatina* complex based on parsimony analysis of 24 individuals sequenced for approximately 700 base pairs (bp) of the cytochrome *b* (cyt *b*) gene. Their best estimate of phylogenetic relationships consisted of two major clades, one comprised of populations from the southern Sierra Nevada (southern *platensis* and *croceater*) plus *klauberi* of the Transverse and Peninsular ranges of southern California, and a second clade made up of populations from the southern and central coast regions (*xanthoptica* and *eschscholtzii*). Both *oregonensis* and *platensis* were recovered as paraphyletic, with three divergent (up to 15% corrected pairwise sequence divergence) lineages within *oregonensis* and a deep north-south break (14%) within *platensis* across the central Sierra Nevada. Samples of *oregonensis* were usually basal, though relationships at the base of the tree lacked resolution.

Kuchta et al. (2009b) sequenced 385 individuals from 224 populations for approximately 450 bp of cyt *b* in a Bayesian phylogenetic analysis to reexamine Stebbins' (1949) historical biogeographic scenario and Moritz et al.'s (1992) mtDNA phylogeography with increased sampling and relaxed molecular clock divergence time estimation. As in Moritz et al. (1992), Kuchta et al. (2009b) found support for a clade made up of populations from the southern Sierra Nevada and mountains of southern California (southern *platensis*, *croceater*, and *klauberi*) which they refer to as the "inland clade", and a clade made up of *eschscholtzii* and *xanthoptica*, referred to as the "coastal clade" (Fig. 2A). A third well-supported clade made up of "*oregonensis*" populations found along northern coastal California was also recovered, and designated

oregonensis [1]. In addition to the coastal, inland, and *oregonensis* [1] clades, Kuchta et al. refer to a more inclusive group comprising the inland clade plus the remainder of the complex as "Clade A" (though as the authors note, this is not a well-supported clade and the term is used only for discussion purposes). Consistent with the results from Moritz et al. (1992), *oregonensis* and *platensis* were both recovered as paraphyletic. The coastal and inland clades were not grouped as sister taxa in the vast majority of their 95% credible set of trees. Similar to the results from Moritz et al. (1992), Kuchta et al.'s (2009b) Bayesian phylogenetic hypothesis lacked resolution at the base of the tree, due to unstable placement of the *oregonensis* [1] lineage (Fig. 2B).

Although useful as a starting point for inferring evolutionary relationships within *Ensatina*, these mtDNA gene trees may be problematic for inferring the historical biogeography of the complex if the mitochondrial genealogical history differs from the true relationships between lineages (the "species" tree). A high degree of discord among gene trees is expected in samples of closely related populations that have diverged recently (Pamilo and Nei 1988), as in the *E. eschscholtzii* complex. Here, we use multiple DNA sequence based markers to reexamine the evolutionary relationships of members of the *E. eschscholtzii* complex. Our primary inference goal is to provide further resolution to the phylogeny, particularly at the base of the tree, and secondarily to examine the relative performance of a supermatrix concatenation approach compared to a species tree approach in a sample of populations that are experiencing gene flow.

Materials and Methods

TAXON SAMPLING

A total of 49 individual *E. eschscholtzii* were included in this study (Table 1). Where possible, we sampled multiple individuals from major mitochondrial clades identified by Moritz et al. (1992) and Kuchta et al. (2009b) (Fig. 2C). An effort was made to sample genetically "pure" populations, identified based on clustering analyses of allozymes (Pereira and Wake 2009), to avoid introgression as a potential sources of gene tree-species tree discordance (Leaché 2009). These populations are largely concordant with the mtDNA clades identified by Kuchta et al. (2009b) (Pereira and Wake 2009). For clarity, hereafter we follow the terminology of Kuchta et al. (2009b) in using "coastal" and "inland" clades, "Clade A", Northern and Southern *platensis*, *oregonensis* [1], [2], [3], and [4], and *xanthoptica* [1] and [2].

MOLECULAR MARKERS

We sequenced one mitochondrial gene (cyt *b*) and six nuclear protein-coding genes, three of which were taken from the literature (*CXCR4*, *RAG1*, *SLC8A3*) (Roelants et al. 2007), and three that we developed from a cDNA library constructed from an individual *E. e. xanthoptica* (*MLC1*, *MLC2*, *RPL12*) (Table 2). Gene fragments sequenced from the cDNA library were identified using the Metazome v2.0.4 project database (www.metazome.net). DNA was extracted from tissues (liver or tail-tip) using Qiagen DNeasy tissue kits following the manufacturer's protocol (Qiagen). PCRs consisted of 40 cycles of 94˚C for 30 s, Ta˚C for 45 s, and 72˚C for 60 s, with locus-specific annealing temperatures (Table 2). PCR products were purified using ethanol

following standard methods. Purified templates were sequenced in both directions on an ABI 3730 capillary sequencer. DNA sequences were edited and aligned using Geneious Pro v4.7 (Drummond et al. 2009). Sequences are deposited in GenBank (Accession Nos. XXXX-XXXX).

CONCATENATED PHYLOGENY

To date, existing species tree methods require the user to assign individuals to species a priori, a difficult problem when the assignment is unknown, as is often the case with recent radiations or in samples of closely related populations (Belfiore et al. 2008; Leaché 2009; Fujita et al. 2010). One approach has been to first use a method such as concatenation (or genetic clustering, network approaches, etc.) to identify lineages, and then estimate their phylogenetic relationships using the multispecies coalescent (Leaché 2009; Fujita et al. 2010). Although concatenation has limitations compared to methods that model the relationship between the species tree and gene trees embedded within them (Degnan and Rosenberg 2009; Edwards 2009; Heled and Drummond 2010), concatenation may still provide a useful starting point for species tree inference (Leaché 2009; Fujita et al. 2010). The mtDNA and six nuclear loci were concatenated to conduct a Bayesian analysis partitioned by gene and by codon in MrBayes v3.1.2 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). The Akaike information criterion (AIC) implemented in MrModeltest v2 (Nylander 2004) was used to choose the best-fit nucleotide substitution model for each gene and codon. We used *Aneides lugubris*, *Hydromantes platycephalus* and *Plethodon vehiculum* as outgroups. Four independent analyses were run for 20,000,000 generations after a burn-in of 1,000,000 generations, each using random starting trees and default priors. In each analysis, four Markov chains (using default heating values) were sampled every 1,000 generations. Stationarity was evaluated by visually examining likelihood and parameter trace files from MCMC runs in Tracer v1.5 (Rambaut and Drummond 2007). Trees sampled prior to reaching stationarity were discarded as burn-in, and the remainder used to construct a 50% majority rule consensus tree. Posterior probability values were used to assess phylogenetic support; clades with values greater than 95% were considered significantly supported.

BAYESIAN INFERENCE OF SPECIES TREES

Because coalescent species-tree methods may be more sensitive to missing data than approaches such as concatenation (Thomson et al. 2008; Edwards 2009), we used a pruned dataset for our species tree analysis. In some cases, we combined data from two individuals of the same lineage (identified based on results from the concatenated phylogeny) to create a chimeric individual with complete data for all loci. This resulted in 10 "species," including representatives of the four major lineages of *oregonensis*, Northern and Southern *platensis*, *croceater*, *klauberi*, *xanthoptica* [2] and *eschscholtzii*. We did not include a representative of *xanthoptica* [1] or *picta* because of too much missing data. We used a Bayesian Markov chain Monte Carlo (MCMC) method under a multispecies coalescent model to estimate the species tree using BEAST v1.5.4 (Drummond and Rambaut 2007; Heled and Drummond 2010). Four independent analyses were run for 100,000,000 generations each after a burn-in of 10,000,000 generations. Adequate mixing of the MCMC was evaluated by examining ESS (expected sample size) trace files for estimated parameters in Tracer v1.5 (Rambaut and Drummond 2007). Trees sampled prior to reaching stationarity were discarded as burn-in, and the remainder were combined across runs to

construct the maximum clade credibility tree using TreeAnnotator v1.5.3 (Drummond and Rambaut 2007).

Results

MOLECULAR MARKER POLYMORPHISM

The aligned sequenced matrix has a length of 3978 bp. We were unable to amplify all loci for every individual, suggesting mutations in primer binding sites. Not surprisingly, the nuclear protein coding loci showed relatively little variation compared to the mtDNA locus. *MLC2* showed the highest level of polymorpism among the nuclear loci, and *SLC8A3* the lowest (Table 2).

CONCATENATED PHYLOGENY

Results from the concatenated Bayesian analysis are shown in Figure 2D. Individuals are labeled according to the distribution of mtDNA clades identified by Kuchta et al. (2009b) for ease of comparison. Basal relationships were not well-resolved. Inland and coastal clades were recovered with strong support. Within the inland clade, *klauberi* forms a well-supported clade sister to southern *platensis* and *croceater*. One individual (sample 39) morphologically assigned to *platensis* groups with the three *croceater* samples with strong support, rendering southern *platensis* paraphyletic with respect to *croceater* (Table 1, Fig. 2D). *Oregonensis* [2] forms a strongly supported clade within a larger clade containing individuals assigned to *picta* and *oregonensis* [3]. *Picta* is paraphyletic with respect to *oregonensis* [2]. *Oregonensis* [4] forms another well-supported clade, but its relationship to other coastal clades is also uncertain. The inland clade consisting of *eschscholtzii* and *xanthoptica* [1] and [2] was recovered with high support, sister to a clade comprising *oregonensis* [1] individuals.

SPECIES TREE

Results from the Bayesian species tree analysis are shown in Figure 3. Most parameters had ESS values greater than 200 indicating good mixing of the MCMC. The ESS of the likelihood parameter was ~1100. Overall, major basal relationships are well-resolved, but support values are low for other nodes. *Oregonensis* [1] is strongly supported as part of the coastal clade containing *xanthoptica* and *eschscholtzii*, though whether it is sister to *xanthoptica* or *eschscholtzii* is uncertain. The remaining individuals (*klauberi*, *croceater*, northern and southern *platensis*, *oregonensis* [2], [3], and [4]) form a clade that is close to being considered well-supported (posterior probability value of 0.93). Relationships within this clade are not well-resolved.

Discussion

CONCATENATED TREE

The concatenated phylogeny shows an overall high degree of similarity to the mtDNA gene genealogies of Mortiz et al. (1992) and Kuchta et al. (2009b). As in those studies, our analyses provided strong support for inland (*klauberi*, southern *platensis*, and *croceater*) and coastal (*eschscholtzii* and *xanthoptica*) clades. Similarly, *oregonensis* and *platensis* were both found to be paraphyletic, rendering our concatenated tree unresolved for basal relationships. Phylogeographic patterns in northern California are complex. Samples from areas in northern California that fall within the range of *oregonensis* [3] as delimited by Kuchta et al. (2009b) are divergent from samples assigned to the same clade from Washington and Oregon. Most individuals assigned to the subspecies *picta* group together, but are nested within *oregonensis*. Sample 11, an individual morphologically classified as *picta* collected from Stebbins' (1949) *picta-oregonensis* intergrade zone, is grouped with samples of *oregonensis* [2] (Fig. 2D). The ranges of *oregonensis* [3] and [4] apparently overlap in Trinity and Shasta counties where the Cascade Range region meets the Klamath Range region to the west: geographically proximate samples 7 and 8 morphologically designated as *oregonensis* are recovered in divergent clades (*oregonenis* [4] and *oregonensis* [3], respectively). Allozymes show the same pattern (Jackman and Wake 1994; Pereira and Wake 2009). A major difference between our results and those from Kuchta et al. (2009) is the placement of *oregonensis* [1], a clade distributed along portions of coastal northern California (Fig. 2C). Kuchta et al. (2009b) were unable to confidently place this lineage, recovering four possible statistically indistinguishable topologies (Fig. 2B). Ultimately, they concluded that *oregonensis* [1] was sister to their "clade A" based on their majority rule consensus topology (Fig. 2A). Our concatenated analyses however, place this lineage sister to the coastal clade with strong support (Fig. 2D). Northern *platensis* forms a well-supported clade, but its relationship to other major clades at the base of the tree is unresolved. Within the inland clade, *klauberi* is strongly supported as sister to *croceater* + southern *platensis*; sample number 39 renders southern *platensis* paraphyletic with respect to *croceater*.

SPECIES TREE

The deepest division in our species tree (Fig. 3) separates a coastal clade, including *oregonensis* [1], from the remainder of the complex with strong support. Within the second, more inclusive clade, support values are much lower compared to the concatenated tree, a result that may be expected for species tree approaches (Thomson et al. 2008; Edwards 2009). The low support for these relationships may be due to violations of model assumptions, namely gene flow (Eckert and Carstens 2008). Genetic admixture of divergent populations within *Ensatina* subspecies is known to be geographically broad (up to 100 km), for example between distinct lineages of *oregonensis* and between northern and southern *platensis* (Pereira and Wake 2009). Despite our efforts to sample genetically "pure" populations away from contact zones, some of our samples are likely individuals with admixed ancestry. Our statistical confidence in the majority of the species tree relationships is low, a result that is predicted for samples of closely related populations that are experiencing gene flow. Despite the low support values, at least some of the relationships seem plausible from a biogeographic standpoint (e.g., *oregonensis* [4] as sister to [2] + [3]), but others are difficult to reconcile without invoking extinction (e.g., the paraphyly of the inland clade with respect to *klauberi*).

LIMITATIONS

Our results highlight the difficult nature of inferring species trees among samples of closely related populations that have diverged recently and are experiencing gene flow. Methodological advances in species tree inference that allow for periods of gene flow during speciation events may provide a solution to this problem as we continue the transition from single gene to multilocus, model-based analyses.

Acknowledgements

We thank Craig Moritz, Jimmy A. McGuire, David B. Wake, and George K. Roderick for comments on an earlier draft of this manuscript. For collecting and donating samples, we thank Brad Alexander, Dave Goodward, Bob Hansen, Brad Hollingsworth, Michelle Koo, Ted Papenfuss, Jim Parham, Jorge Valdez-Villavicencio, Jens Vindum, and Anny Peralta-Garcia. Funding was provided by the Museum of Vertebrate Zoology Martens and Louise Kellogg funds and the National Science Foundation (Doctoral Dissertation Improvement Grant DEB-0909821 and NSF DEB-0641078). We thank the ESRI Conservation Program for providing ArcGIS software for our use.

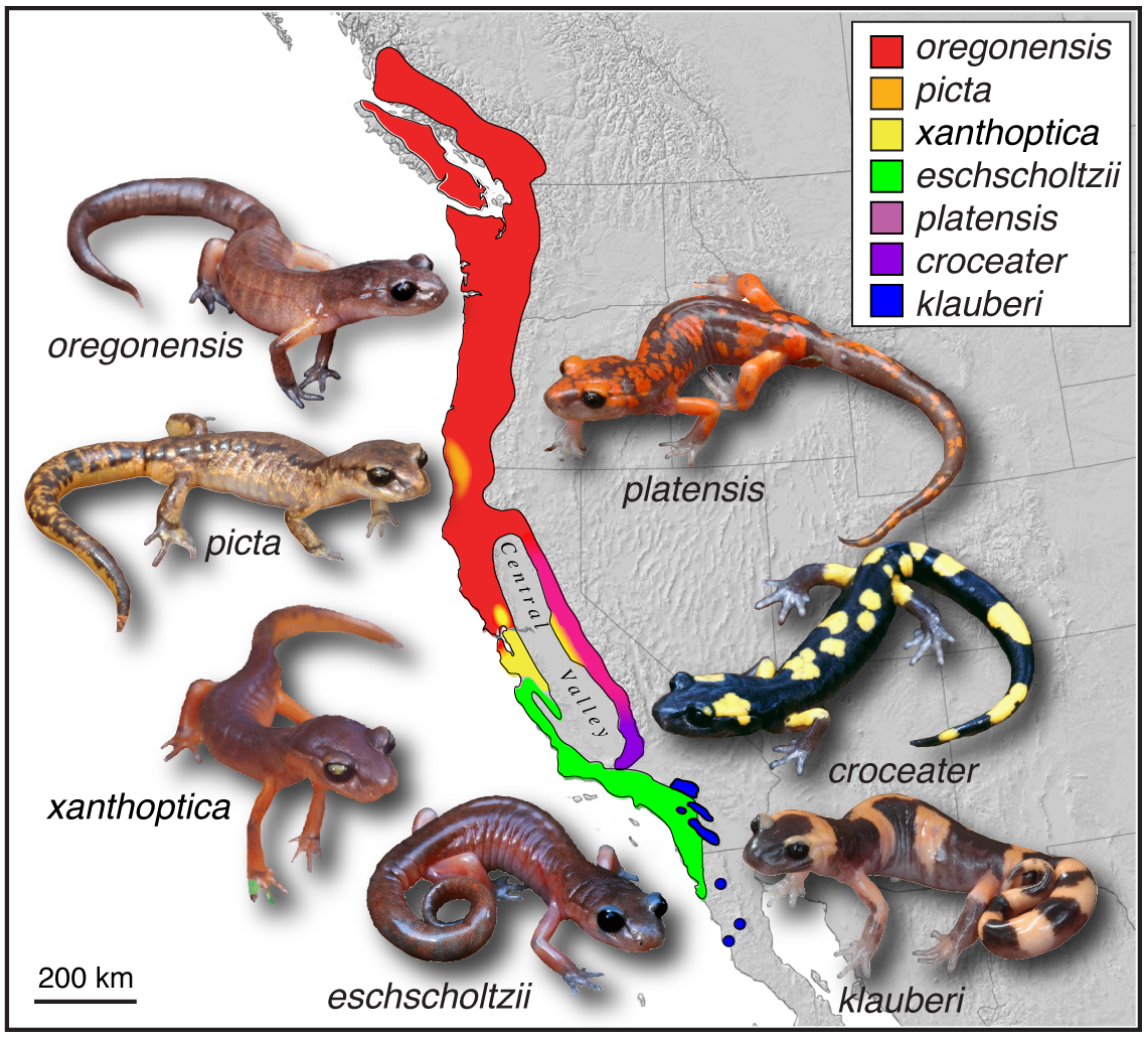

Figure 1. Distribution of the *Ensatina eschscholtzii* complex showing the ranges of the seven currently recognized subspecies.

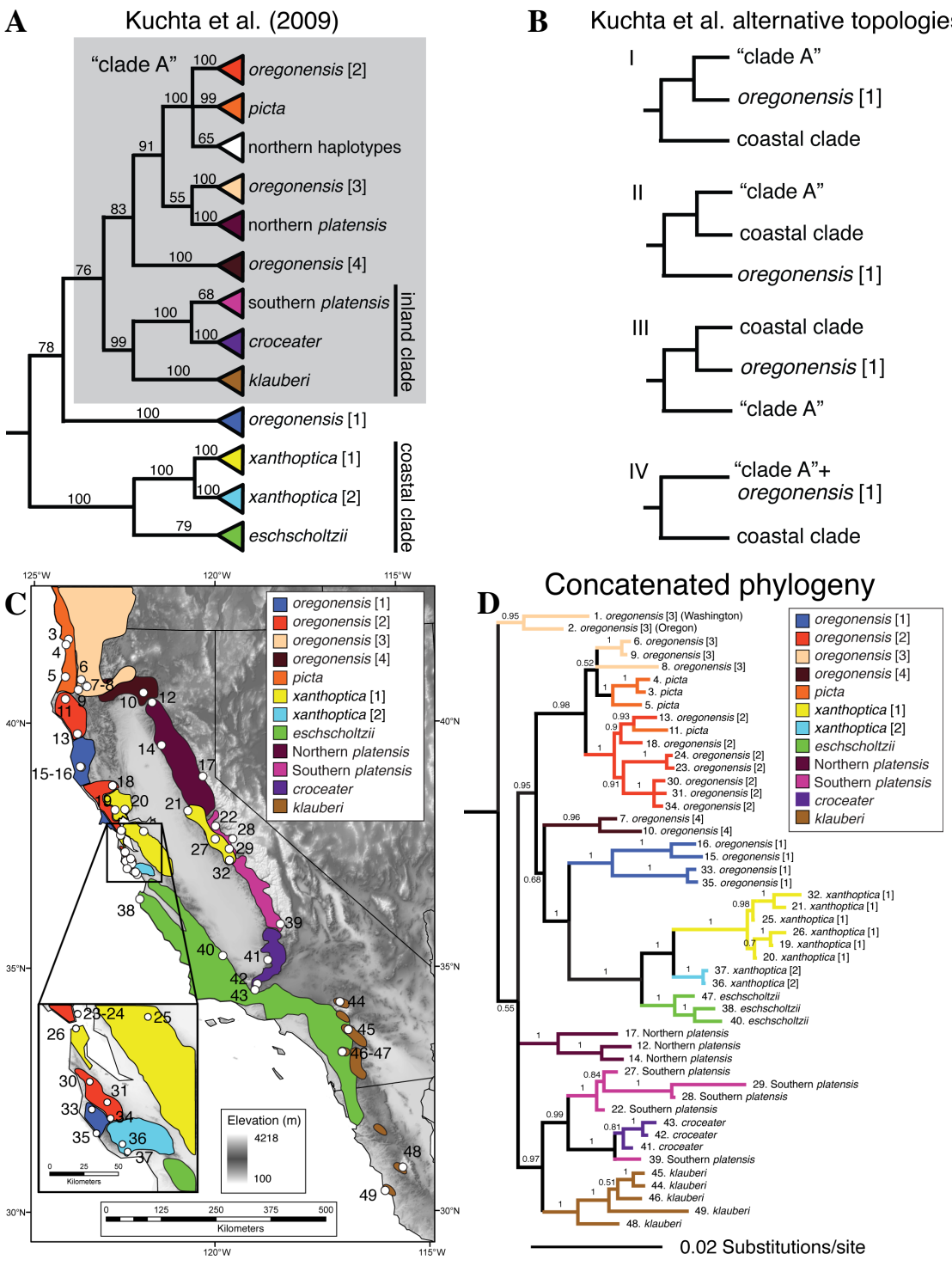

Figure 2. A) majority rule Bayesian mtDNA genealogy from Kuchta et al. (2009); B) alternative, statistically indistinguishable topologies from Kuchta et al. (2009); C) sampling superimposed on mtDNA clades identified by Kuchta et al. (2009) (samples 1 and 2 from Washington and Oregon not shown); D) multilocus phylogeny of concatenated data from this study. Numbers above branches in A and D represent posterior probability values.

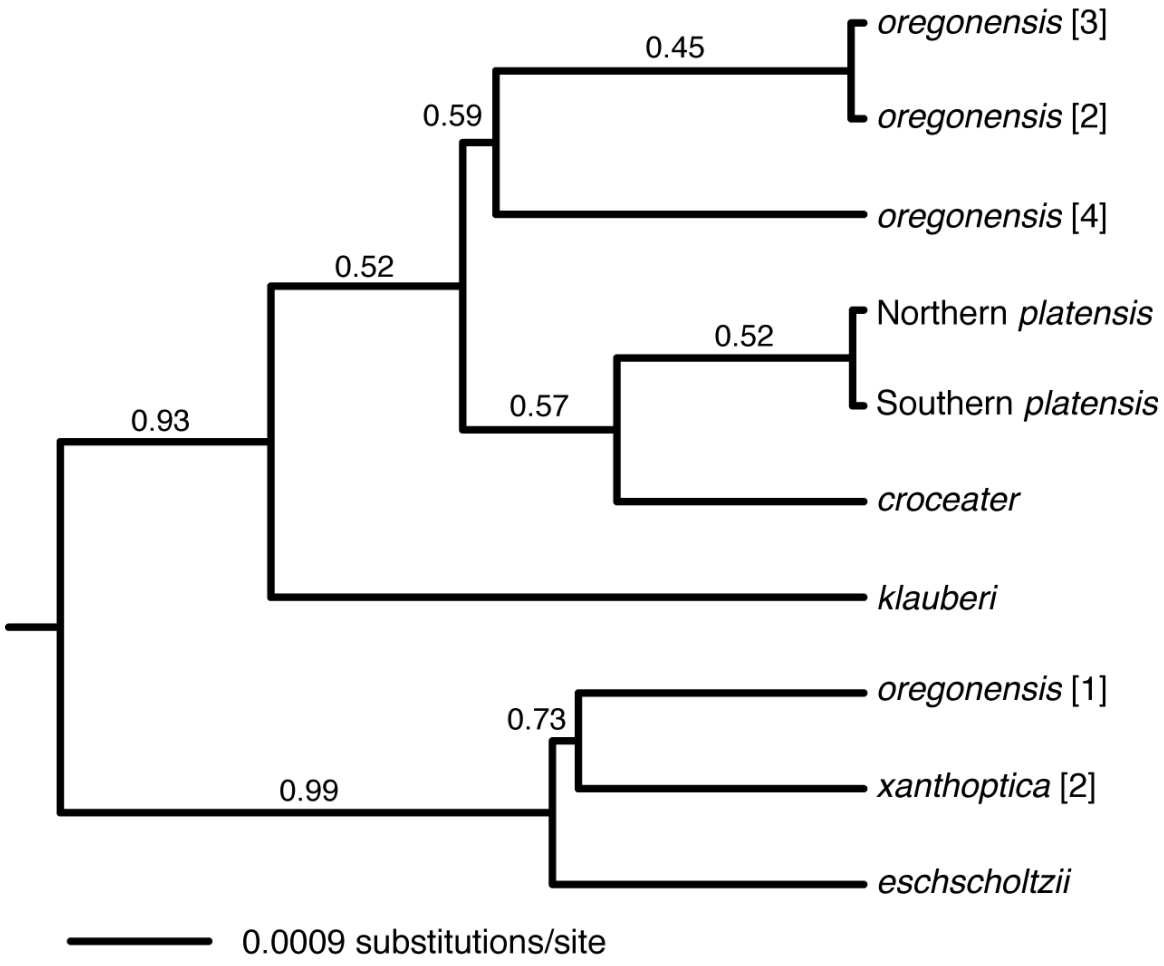

Figure 3. Species tree from BEAST analysis. Numbers above branches represent posterior probability values.

Table 1. Taxon sampling for phylogenetic analyses. Map # corresponds to number on Fig. 2C. Subspecies assignment is based on color pattern and distribution determined by original collector as listed in the MVZ database. MtDNA clade designations are expectations only based on the distribution of clades defined by Kuchta et al. (2009).

Map #	ID Number	Subspecies	MtDNA clade	Latitude	Longitude	State	County
1	MVZ 230987	*oregonensis*	*oregonensis* [3]	47.15450	-121.07830	Washington	Kittitas
2	MVZ 222572	*oregonensis*	*oregonensis* [3]	44.57440	-122.45740	Oregon	Linn
3	MVZ 220614	*picta*	*picta*	41.75281	-124.01813	California	Del Norte
4	DBW 6482	*picta*	*picta*	41.64415	-124.08176	California	Del Norte
5	RP 209	*picta*	*picta*	40.98335	-124.05930	California	Humboldt
6	RP 217	*oregonensis*	*oregonensis* [3]	40.94155	-123.63504	California	Humboldt
7	MVZ 237551	*oregonensis*	*oregonensis* [4]	40.82591	-123.51317	California	Trinity
8	RP 212	*oregonensis*	*oregonensis* [3]	40.80391	-123.47127	California	Trinity
9	MVZ 211881	*oregonensis*	*oregonensis* [3]	40.73061	-123.69726	California	Humboldt
10	MVZ 197518	*oregonensis*	*oregonensis* [4]	40.71295	-121.94613	California	Shasta
11	MVZ 220589	*picta*	*oregonensis* [2]	40.52858	-124.03473	California	Humboldt
12	MVZ 234865	*platensis*	Northern *platensis*	40.51857	-121.70877	California	Shasta
13	DBW 6578	*oregonensis*	*oregonensis* [2]	39.83195	-123.67661	California	Mendocino
14	MVZ 225028	*platensis*	Northern *platensis*	39.64692	-121.43832	California	Butte
15	CM 140	*oregonensis*	*oregonensis* [1]	39.16212	-123.56869	California	Mendocino
16	MVZ 194896	*oregonensis*	*oregonensis* [1]	39.15260	-123.54166	California	Mendocino
17	MVZ 225030	*platensis*	Northern *platensis*	39.01371	-120.33931	California	El Dorado
18	MVZ 223052	*oregonensis*	*oregonensis* [2]	38.79681	-122.70224	California	Lake
19	MVZ 237582	*xanthoptica*	*xanthoptica* [1]	38.30214	-122.62262	California	Sonoma
20	MVZ 238078	*xanthoptica*	*xanthoptica* [1]	38.31588	-122.36464	California	Napa
21	MVZ 243185	*xanthoptica*	*xanthoptica* [1]	38.31032	-120.72275	California	Calaveras
22	TJD 303	*platensis*	Southern *platensis*	38.00083	-120.00500	California	Tuolumne
23	MVZ 237491	*oregonensis*	*oregonensis* [2]	37.88855	-122.44842	California	Marin
24	MVZ 237492	*oregonensis*	*oregonensis* [2]	37.88855	-122.44842	California	Marin
25	MVZ 243244	*xanthoptica*	*xanthoptica* [1]	37.87945	-121.86513	California	Contra Costa
26	CAS 231482	*xanthoptica*	*xanthoptica* [1]	37.79231	-122.45972	California	San Francisco
27	TJD 182	*platensis*	Southern *platensis*	37.73583	-120.01047	California	Mariposa
28	MVZ 249898	*platensis*	Southern *platensis*	37.74059	-119.55998	California	Mariposa

29	MVZ 243315	*platensis*	Southern *platensis*	37.53976	-119.64956	California	Mariposa
30	MVZ 222997	*oregonensis*	*oregonensis* [2]	37.44826	-122.33497	California	San Mateo
31	MVZ 222987	*oregonensis*	*oregonensis* [2]	37.31510	-122.18600	California	San Mateo
32	MVZ 244091	*xanthoptica*	*xanthoptica* [1]	37.30500	-119.64740	California	Madera
33	MVZ 244308	*oregonensis*	*oregonensis* [1]	37.26610	-122.31230	California	San Mateo
34	MVZ 247636	*xanthoptica*	*oregonensis* [2]	37.21200	-122.15600	California	Santa Cruz
35	MVZ 230790	*oregonensis*	*oregonensis* [1]	37.11167	-122.26970	California	Santa Cruz
36	MVZ 230886	*xanthoptica*	*xanthoptica* [2]	37.04444	-122.05389	California	Santa Cruz
37	MVZ 230847	*xanthoptica*	*xanthoptica* [2]	36.99555	-122.01055	California	Santa Cruz
38	MVZ 239635	*eschscholtzii*	*eschscholtzii*	36.49279	-121.93270	California	Monterey
39	MVZ 244154	*platensis*	Southern *platensis*	35.99100	-118.36567	California	Tulare
40	SSSweet 32453	*eschscholtzii*	*eschscholtzii*	35.35357	-119.82162	California	Kern
41	TJD 307	*croceater*	*croceater*	35.26444	-118.69361	California	Kern
42	MVZ 234859	*croceater*	*croceater*	34.77470	-118.96940	California	Ventura
43	MVZ 195607	*croceater*	*croceater*	34.65289	-119.02541	California	Ventura
44	JFP 517	*klauberi*	*klauberi*	34.37198	-116.92859	California	San Bernardino
45	TJP 29472A	*klauberi*	*klauberi*	33.79685	-116.74747	California	Riverside
46	CM 56	*klauberi*	*klauberi*	33.34615	-116.91318	California	San Diego
47	TJD 295	*eschscholtzii*	*eschscholtzii*	33.34424	-116.91919	California	San Diego
48	MVZ 229220	*klauberi*	*klauberi*	30.93333	-115.56667	Baja California	
49	SDField 573	*klauberi*	*klauberi*	30.46853	-116.00833	Baja California	

Table 2. PCR annealing temperatures, approximate sequence length, primer sequences, and average percent pairwise sequence identity for a given alignment.

Locus[1]	Size (bp)	Annealing temp.	PCR Primer[2]	Primer sequence (5´-3´)	% Pairwise Identity
Cyt *b*	804	48˚C	MVZ 15	GAACTAATGGCCCACACWWTACGNAA	89.9%
			MVZ 16	AAATAGGAARTATCAYTCTGGTTTRAT	
CXCR4	405	55˚C	CXCR4F	GTGGTCTGTGGATGCTGTCAT	98.4%
			CXCR4R	TGCAGTAGCAGATCAAGATGA	
SLC8A3	955	58˚C	SLC8A3F	CATTCGGGTCTGGAATGAAA	99.0%
			SLC8A3R	ACACCACCATCCCCTCTGTA	
RAG1	885	56˚C	RAG1F1	ACAGGATATGATGARAAGCTTGT	98.7%
			RAG1R	TTRGATGTGTAGAGCCAGTGGTGYTT	
MLC1	331	55˚C	C71	TAATTGATACAGCCATTGGAGTCTTC	96.0%
			C72	AGTATAACCAGTGCGGTGATGTTATG	
MLC2	293	66˚C	C3	ATGCGTGTGAATTCCACATAATTG	98.0%
			C4	GAAGAACCCAACTGATGAATACCT	
RPL12	305	65˚C	230AF	ATCGTCAATGATGTCATGTGGT	97.8%
			230CR	AACTGGTGACTGGAAGGGACT	

[1]Cytochrome *b*, CXC chemokine receptor 4, Solute carrier family 8, Recombination activating gene 1, Myosin regulatory light chain 1, Myosin regulatory light chain 2, Ribosomal protein L12.

[2]PCR primer references: cyt *b* (Moritz et al. 1992); CXCR4, SLC8A3, RAG1 (Roelants et al. 2007); MLC1, MLC2, RPL12 (this study).

Literature Cited

Alberto, F. 2009. MsatAllele_1.0: An R Package to Visualize the Binning of Microsatellite Alleles. J Hered 100:394-397.

Alexandrino, J., S. J. E. Baird, L. Lawson, J. R. Macey, C. Moritz, and D. B. Wake. 2005. Strong selection against hybrids at a hybrid zone in the *Ensatina* ring species complex and its evolutionary implications. Evolution 59:1334-1347.

Anderson, E. C. 2008. Bayesian inference of species hybrids using multilocus dominant genetic markers. Philosophical Transactions of the Royal Society B 363:2841-2850.

Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217-1229.

Andersson, M. 1994. Sexual Selection. Princeton University Press, Princeton.

Araújo, M. B., and R. G. Pearson. 2005. Equilibrium of species' distributions with climate. Ecography 28:693-695.

Arévalo, E., S. K. Davis, and J. W. Sites, Jr. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Systematic Biology 43:387-418.

Armsworth, P. R., and J. E. Roughgarden. 2008. The structure of clines with fitness-dependent dispersal. The American Naturalist 172:648-657.

Arnold, J. 1993. Cytonuclear disequilibria in hybrid zones. Annual Review of Ecology and Systematics 24:521-524.

Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling 157:101-118.

Avise, J. A., and N. C. Saunders. 1984. Hybridization and introgression among species of sunfish (*Lepomis*): analysis by mitochondrial DNA and allozyme markers. Genetics 108:237-255.

Axelrod, D. I. 1966. The Pleistocene Soboba flora of southern California. University of California Publications in Geological Sciences 60:1-79.

Axelrod, D. I. 1981. Holocene Climatic Changes in Relation to Vegetation Disjunction and Speciation. The American Naturalist 117:847-870.

Axelrod, D. I. 1989. Age and origin of chaparral. Pp. 1-19 *in* S. C. Keely, ed. The California Chaparral: Paradigms Reexamined, Los Angeles.

Bartlein, P. J., K. H. Anderson, P. M. Anderson, M. E. Edwards, C. J. Mock, R. S. Thompson, R. S. Webb, T. Webb III, and C. Whitlock. 1998. Paleoclimate simulations for North America over the past 21,000 years: features of the simulated climate and comparisons with paleoenvironmental data. Quaternary Science Reviews 17:549-585.

Barton, N. H. 2000. Genetic hitchhiking. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 355:1553-1562.

Barton, N. H., and S. J. E. Baird. 1995. Analyse - An application for analyzing hybrid zones. Freeware, Edinburgh.

Barton, N. H., and K. Gale. 1993. Genetic analysis of hybrid zones. Pp. 13-45 *in* R. G. Harrison, ed. Hybrid Zones and the Evolutionary Process. Oxford University Press, Oxford.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual Review of Ecology and Systematics 16:113-148.

Belfiore, N. M., L. Liu, and C. Moritz. 2008. Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: Geomyidae). Systematic Biology 57:294-310.

Bell, R., J. L. Parra, M. Tonione, C. Hoskin, J. MacKenzie, S. Williams, and C. Moritz. 2010. Patterns of persistence and isolation indicate resilience to climate change in montane rainforest lizards. Molecular Ecology in press.
Boecklen, W. J., and D. J. Howard. 1997. Genetic analysis of hybrid zones: number of markers and power of resolution. Ecology 78:2611-2616.
Bridle, J. R., S. J. E. Baird, and R. K. Butlin. 2001. Spatial structure and habitat variation in a grasshopper hybrid zone. Evolution 55:1832 - 1843.
Brito, P. H., and S. V. Edwards. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica 135:439-455.
Brown, C. W. 1974. Hybridization among the subspecies of the plethodontid salamander *Ensatina eschscholtzii*. University of California Publications in Zoology 98:1-57.
Brown, C. W., and R. C. Stebbins. 1964. Evidence for hybridization between the blotched and unblotched subspecies of the salamander *Ensatina eschscholtzii*. Evolution 18:706-707.
Brunk, H. D. 1955. Maximum likelihood estimates of monotone parameters. Ann Math Stat 26:607 - 616.
Cain, A. J. 1954. Animal Species and Their Evolution. Hutchinson House, London.
Calinski, R. B., and J. Harabasz. 1974. A dendrite method for cluster analysis. Communications in Statistics 3:1-27.
Carnaval, A. C., M. J. Hickerson, C. F. B. Haddad, M. T. Rodrigues, and C. Moritz. 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. Science 323:785-789.
Carstens, B. C., and C. L. Richards. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. Evolution 61:1439-1454.
Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA.
Davic, R. D., and H. H. Welsh, Jr. 2004. On the ecological roles of salamanders. Annual Review of Ecology and Systematics 35:405-434.
Degnan, J. H., and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution 24:332-340.
Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 39:1-38.
DeutschesKlimarechenzentrumModellbetreuungsgruppe. 1992. The ECHAM3 atmospheric general circulation model: DKRZ Technical Report 6. Deutsches Klimarechenzentrum, Hamburg, Germany.
Devitt, T. J., R. Pereira, L. Jakkula, J. Alexandrino, C. Bardeleben, and C. Moritz. 2009. Isolation and characterization of 15 polymorphic microsatellites in the Plethodontid salamander *Ensatina eschscholtzii*. Molecular Ecology Resources 9:966-969.
Dobzhansky, T. 1958. Species after Darwin. Pp. 19-55 *in* S. A. Barnett, ed. A Century of Darwin. Heinemann, London.
Dobzhansky, T., L. Ehrman, O. Pavlovsky, and B. Spassky. 1964. The superspecies *Drosophila paulistorum*. Proceedings of the National Academy of Sciences USA 51:3-9.
Dobzhansky, T., and O. Pavlovsky. 1967. Experiments on the incipient species of the *Drosophila paulistorum* complex. Genetics 55:141-156.
Dobzhansky, T., and B. Spassky. 1959. *Drosophila paulistorum*, a cluster of species *in status nascendi*. Proceedings of the National Academy of Sciences USA 45:419-428.

Drummond, A. J., B. Ashton, M. Cheung, J. Heled, M. Kearse, R. Moir, S. Stones-Havas, T. Thierer, and A. Wilson. 2009. Geneious.
Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.
Earl, D. A. 2009. Structure Harvester v0.3, from website: http://users.soe.ucsc.edu/~dearl/software/struct_harvest/.
Eckert, A. J., and B. C. Carstens. 2008. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. Molecular Phylogenetics and Evolution 49:832-842.
Edelaar, P., A. M. Siepielski, and J. Clobert. 2008. Matching habitat choice causes directed gene flow: a neglected dimension in evolution and ecology. Evolution 62:2462-2472.
Edwards, A. 1992. Likelihood. Johns Hopkins University Press, Baltimore, MD.
Edwards, S. V. 2009. Is a new and general theory of molecular systematics emerging? Evolution 63:1-19.
Ehrlich, P. R., and P. H. Raven. 1969. Differentiation of populations. Science 165:1228-1232.
Ehrman, L., and M. Wasserman. 1987. The Significance of Asymmetrical Sexual Isolation. Pp. 1-20 *in* M. K. Hecht, B. Wallace, and G. T. Prance, eds. Evolutionary Biology. Plenum Press, New York.
Endler, J. A. 1973. Gene flow and population differentiation. Science 179:243-250.
Endler, J. A. 1977. Geographic variation, speciation, and clines. Princeton University Press, Princeton, N. J.
Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology 14:2611-2620.
Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources in press.
Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. Genetics 131:479-491.
Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.
Falush, D., M. Stephens, and J. K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7:574-578.
Feder, M. E. 1983. Integrating the ecology and physiology of plethodontid salamanders. Herpetologica 39:291-310.
Feng, X., and S. Epstein. 1994. Climatic implications of an 8000-year hydrogen isotope time series from Bristlecone pine trees. Science 265:1079-1081.
Ferrusquía-Villafranca, I. 1993. Geology of Mexico: A Synopsis. Pp. 3-107 *in* T. P. Ramamoorthy, R. Bye, A. Lot, and J. Fa, eds. Biological Diversity of Mexico: Origins and Distribution. Oxford University Press, New York.
Fisher, R. A. 1950. Gene frequencies in a cline determined by selection and diffusion. Biometrics 6:353-361.
Flot, J.-F. 2009. seqphase: a web tool for interconverting phase input/output files and fasta sequence alignments. Molecular Ecology Resources 9999.

Frost, D. R., and D. M. Hillis. 1990. Species in concept and practice: herpetological applications. Herpetologica 46.
Fujita, M. K., J. A. McGuire, S. C. Donnellan, and C. Moritz. 2010. Diversification and persistence at the arid-monsoonal interface: Australia-wide biogeography of the Bynoe's Gecko (*Heteronotia binoei*; Gekkonidae). Evolution 9999.
Futuyma, D. J. 1998. Evolutionary Biology. Sinauer Associates, Inc., Sunderland, MA.
Gavrilets, S. 2003. Models of speciation: what have we learned in 40 years? Evolution 57:2197-2215.
Graham, C. H., C. Moritz, and S. E. Williams. 2006. Habitat history improves prediction of biodiversity in rainforest fauna. Proceedings of the National Academy of Sciences of the United States of America 103:632-636.
Graybeal, A. 1995. Naming species. Systematic Biology 44:237-250.
Guo, S., and E. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.
Haldane, J. B. S. 1948. The theory of a cline. Journal of Genetics 48:277-284.
Hall, C. A. 2007. Introduction to the Geology of Southern California and Its Native Plants. University of California Press, Berkeley.
Hanley, J. A., and B. J. McNeil. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29-36.
Harrison, R. G. 1983. Barriers to gene exchange in closely related cricket species. I. Laboratory hybridization studies. Evolution 37:245-251.
Harrison, R. G. 1990. Hybrid zones: windows on evolutionary process. Oxford surveys in Evolutionary Biology 7:69-128.
Harrison, R. G., and D. M. Rand. 1989. Mosaic hybrid zones and the nature of species boundaries. Pp. 111-133 *in* D. Otte, and J. A. Endler, eds. Speciation and Its Consequences. Sinauer Associates, Sunderland, MA.
Heim, C. D., B. Alexander, R. W. Hansen, J. H. Valdez-Villavicencio, T. J. Devitt, B. D. Hollingsworth, J. A. Soto-Centeno, and C. R. Mahrdt. 2005. Geographic Distribution: Ensatina eschscholtzii klauberi. . Herpetological Review 36:330-331.
Heled, J., and A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution 27:570-580.
Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society 58:247-276.
Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907-913.
Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 359:183-195.
Highton, R. 1998. Is *Ensatina eschscholtzii* a ring-species? Herpetologica 54:254-278.
Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965-1978.
Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, NJ.
Hill, W. G. 1974. Estimation of linkage disequilibrium in randomly mating populations. Heredity 33:229-239.

Hoskin, C. J., M. Higgie, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. Nature 437:1353-1356.
Howard, D. J. 1986. A zone of overlap and hybridization between two ground cricket species. Evolution 40:34-43.
Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9:1322-1332.
Huelsenbeck, J. P., and P. Andolfatto. 2007. Inference of population structure under a Dirichlet process model. Genetics 175:1787-1802.
Huelsenbeck, J. P., and F. Ronquist. 2001a. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-755.
Huelsenbeck, J. P., and F. Ronquist. 2001b. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-755.
Huesenbeck, J. P. a. K. A. C. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annual Review of Ecology and Systematics 28:437-466.
Hugall, A., C. Moritz, A. Moussalli, and J. Stanisic. 2002. Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail Gnarosophia bellendenkerensis (Brazier 1875). Proceedings of the National Academy of Sciences of the United States of America 99:6112-6117.
Huxley, J. S. 1942. Evolution: the Modern Synthesis. George Allen and Unwin, Ltd., London.
Irwin, D. E. 2000. Song variation in an avian ring species. Evolution 54:998-1010.
Irwin, D. E., and J. H. Irwin. 2002. Circular overlaps: rare demonstrations of speciation. The Auk 119:596-602.
Irwin, D. E., J. H. Irwin, and T. D. Price. 2001. Ring species as bridges between microevolution and speciation. Genetica 112-113:223-243.
Jackman, T. R., and D. B. Wake. 1994. Evolutionary and historical analysis of protein variation in the blotched forms of salamanders of the *Ensatina* complex (Amphibia: Plethodontidae). Evolution 48:876-897.
Jordan, D. S. 1905. The origin of species through isolation. Science 22:545-562.
Kaneshiro, K. Y., and L. V. Giddings. 1987. The Significance of Asymmetrical Sexual Isolation and the Formation of New Species. Pp. 29-43 *in* M. K. Hecht, B. Wallace, and G. T. Prance, eds. Evolutionary Biology. Plenum Press, New York.
Kingman, J. F. C. 1982. On the genealogy of large populations. Journal of Applied Probability 19:27-43.
Kingman, J. F. C. 2000. Origins of the coalescent: 1974-1982. Genetics 156:1461-1463.
Knowles, L. L. 2009. Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. Systematic Biology 58:463-467.
Knowles, L. L., B. C. Carstens, and M. L. Keat. 2007. Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. Current Biology 17:940-946.
Kuchta, S. R., A. H. Krakauer, and B. Sinervo. 2008. Why does the Yellow-eyed Ensatina have yellow eyes? Batesian mimicry of Pacific newts (genus *Taricha*) by the salamander *Ensatina eschscholtzii xanthoptica*. Evolution 62:984-990.
Kuchta, S. R., D. Parks, and D. B. Wake. 2009a. Pronounced phylogeographic structure on a small spatial scale: geomorphological evolution and lineage history in the salamander

ring species *Ensatina eschscholtzii* in central coastal California. Molecular Phylogenetics and Evolution 50:240-255.
Kuchta, S. R., D. S. Parks, R. L. Mueller, and D. B. Wake. 2009b. Closing the ring: historical biogeography of the salamander ring species *Ensatina eschscholtzii*. Journal of Biogeography in press.
Lande, R. 1981. Models of speciation by sexual selection on polygenic traits. PNAS 78:3721-3725.
Lauri, R. K. 2004. A Comparative Floristic Study of Palomar Mountain State Park. Pp. i-x, 1-98. Biology. San Diego State University, San Diego.
Leaché, A. D. 2009. Species tree discordance traces to phylogeographic clade boundaries in North American fence lizards (*Sceloporus*). Systematic Biology 58:547-559.
Levene, H. 1949. On a matching problem arising in genetics. The Annals of Mathematical Statistics 20:91-94.
Lowe, J. 2001. Population Structure and Habitat Relationships of Sympatric Populations of the Terrestrial Salamanders *Plethodon elongatus* and *Ensatina eschscholtzii* Before a Prescribed Fire. Pp. 77. Biology. Humboldt State University, Arcata.
Lushai, G., J. A. Allen, D. Goulson, N. MacLean, and D. A. S. Smith. 2005. The butterfly *Danaus chrysippus* (L.) in East Africa comprises polyphyletic, sympatric lineages that are, despite behavioural isolation, driven to hybridization by female-biased sex ratios. Biological Journal of the Linnean Society 86:117-131.
M'Gonigle, L. K., and R. G. FitzJohn. 2010. ASSORTATIVE MATING AND SPATIAL STRUCTURE IN HYBRID ZONES. Evolution 64:444-455.
MacCallum, C. J., B. Nürnberger, N. H. Barton, and J. M. Szymura. 1998. Habitat preference in the *Bombina* hybrid zone in Croatia. Evolution 52:227-239.
Macholán, M., S. J. E. Baird, P. Munclinger, P. Dufková, B. Bímova, and J. Piálek. 2008. Genetic conflict outweighs heterogametic incompatibility in the mouse hybrid zone? BMC Evolutionary Biology 8:271.
Maddison, W. P. 1997. Gene trees in species trees. Systematic Biology 46:523-536.
Mahrdt, C. R., R. H. McPeak, and L. L. Grismer. 1998. The Discovery of *Ensatina eschscholtzii klauberi* (Plethodontidae) in the Sierra San Pedro Mártir, Baja California, México. Herpetological Natural History 6:73-76.
Mallet, J. 2005. Hybridization as an invasion of the genome. Trends in Ecology and Evolution 20:229-237.
Marshall, J. C., and J. W. Sites. 2001. A comparison of nuclear and mitochondrial cline shapes in a hybrid zone in the *Sceloporus grammicus* complex (Squamata: Phrynosomatidae). Molecular Ecology 10:435-449.
Marshall, J. L., M. L. Arnold, and D. J. Howard. 2002. Reinforcement: the road not taken. Trends in Ecology & Evolution 17:558-563.
May, R. M., J. A. Endler, and R. E. McMurtrie. 1975. Gene frequency clines in the presence of selection opposed by gene flow. The American Naturalist 109:659-676.
Mayr, E. 1942. Systematics and the Origin of Species. Dover Publications, New York.
Mayr, E. 1970. Populations, Species, and Evolution: An Abridgement of Animal Species and Evolution. Belknap Press of Harvard University Press, Cambridge, MA.
McCormack, J. E., H. Huang, and L. L. Knowles. 2009. Maximum-likelihood estimates of species trees: how accuracy of phylogenetic inference depends upon the divergence history and sampling design. Systematic Biology 58:501-508.

McPeek, M. A., and S. Gavrilets. 2006. The evolution of female mating preferences: Differentiation from species with promiscuous males can promote speciation. Evolution:1967-1980.
McRae, B. H. 2006. ISOLATION BY RESISTANCE. Evolution 60:1551-1561.
McRae, B. H., and P. Beier. 2007. Circuit theory predicts gene flow in plant and animal populations. Proceedings of the National Academy of Sciences 104:19885-19890.
McRae, B. H., B. G. Dickson, T. H. Keitt, and V. B. Shah. 2008. USING CIRCUIT THEORY TO MODEL CONNECTIVITY IN ECOLOGY, EVOLUTION, AND CONSERVATION. Ecology 89:2712-2724.
Meirmans, P. G. 2010. GenoDive Help files.
Meirmans, P. G., and P. H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Molecular Ecology Notes 4:792-794.
Minnich, R. A. 2007. California climate, paleoclimate and paleovegetation. Pp. 43-70 *in* M. G. Barbour, T. Keeler-Wolf, and A. A. Schoenherr, eds. Terrestrial Vegetation of California. University of California Press, Berkeley.
Minnich, R. A., and E. Franco-Vizcaíno. 2005. Baja California's Enduring Mediterranean Vegetation: Early Accounts, Human Impacts, and Conservation Status. Pp. 370-386 *in* J.-L. E. Cartron, and G. Ceballos, eds. Biodiversity, Ecosystems, and Conservation in Northern Mexico. Oxford University Press, New York.
Moritz, C., C. Schneider, and D. B. Wake. 1992. Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. Systematic Biology 41:273-291.
Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853-858.
Neigel, J. E. 2002. Is FST obsolete? Conservation Genetics 3:167-173.
Nichols, R. 2001. Gene trees and species trees are not the same. Trends in Ecology & Evolution 16:358-364.
Nylander, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
Owen, L. A., R. C. Finkel, R. A. Minnich, and A. E. Perez. 2003. Extreme southwestern margin of late Quaternary glaciation in North America: Timing and controls. Geology 31:729-732.
Pamilo, P., and M. Nei. 1988. Relationships between gene trees and species trees. Molecular Biology and Evolution 5:568-583.
Pearson, R. G., W. Thuiller, M. B. Araújo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, T. P. Dawson, and D. C. Lees. 2006. Model-based uncertainty in species range prediction. Journal of Biogeography 33:1704-1711.
Pella, J., and M. Masuda. 2006. The Gibbs and split-merge sampler for population mixture analysis from genetic data with incomplete baselines. Canadian Journal of Fisheries & Aquatic Sciences 63:576-596.
Pereira, R., and D. B. Wake. 2009. Genetic leakage after adaptive and non-adaptive divergence in the *Ensatina eschscholtzii* ring species. Evolution 63:2288-2301.
Peterson, A. T., and Á. S. Nyári. 2008. Ecological niche conservatism and Pleistocene refugia in the Thrush-like Mourner (*Schiffornis sp.*) in the Neotropics. Evolution 62:173-183.

Phillips, B. L., S. J. E. Baird, and C. Moritz. 2004. When vicars meet: A narrow contact zone between phylogeographic lineages of the rainforest skink, *Carlia rubrigularis*. Evolution 58:1536-1548.
Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190.
Pialek, J., and N. H. Barton. 1997. The Spread of an Advantageous Allele Across a Barrier: The Effects of Random Drift and Selection Against Heterozygotes. Genetics 145:493-504.
Porter, A. H., R. Wenger, H. Geiger, A. Scholl, and A. M. Shapiro. 1997. The *Pontia daplidice-edusa* Hybrid Zone in Northwestern Italy. Evolution 51:1561-1573.
Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.
Rambaut, A., and A. J. Drummond. 2007. Tracer.
Randler, C. 2002. Avian hybridization, mixed pairing and female choice. Animal Behaviour 63:103-119.
Raymond, M., and F. Rousset. 1995. GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248-249.
Richards, C. L., B. C. Carstens, and L. Lacey Knowles. 2007. Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. Journal of Biogeography 34:1833-1845.
Roelants, K., D. J. Gower, M. Wilkinson, S. P. Loader, S. D. Biju, K. Guillaume, L. Moriau, and F. Bossuyt. 2007. Global patterns of diversification in the history of modern amphibians. Proceedings of the National Academy of Sciences 104:887-892.
Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.
Rosenberg, N. A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4:137-138.
Rousset, F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8:103-106.
Schaffer, J. P. 1993. California's Geological History and Changing Landscapes. Pp. 49-54 *in* J. C. Hickman, ed. The Jepson Manual: Higher Plants of California. University of California Press, Berkeley.
Schluter, D. 2009. Evidence for ecological speciation and its alternative. Science 323:737-741.
Sites, J. W., N. H. Barton, and K. M. Reed. 1995. The genetic structure of a hybrid zone between two chromosome races of the *Sceloporus grammicus* complex (Sauria, Phrynosomatidae) in Central Mexico. Evolution 49:9-36.
Slatkin, M. 1973. Gene flow and selection in a cline. Genetics 75:733-756.
Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science 236:787-792.
Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic Zoology 35:627-632.
Spaulding, W. G. 1990. Vegetational and climatic development of the Mojave Desert: The last glacial maximum to the present *in* J. L. Betancourt, T. R. Van Devender, and P. A. Martin, eds. Packrat Middens: The last 40,000 years of biotic change. University of Arizona Press, Tucson.

Staub, N. L., C. W. Brown, and D. B. Wake. 1995. Patterns of growth and movements in a population of *Ensatina eschscholtzii platensis* (Caudata: Plethodontidae) in the Sierra Nevada, California. Journal of Herpetology 29:593-599.
Stebbins, R. C. 1949. Speciation in Salamanders of the Plethodontid Genus *Ensatina*. University of California Publications in Zoology 54:47-124.
Stebbins, R. C. 1954. Natural history of the salamanders of the plethodontid genus *Ensatina*. University of California Publications in Zoology 54:47-124.
Stebbins, R. C. 1957. Intraspecific sympatry in the lungless salamander *Ensatina eschscholtzii*. Evolution 11:265-270.
Stebbins, R. C. 2003. A Field Guide to Western Reptiles and Amphibians. Houghton Mifflin Company, New York.
Stephens, M., and P. Donnelly. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. American Journal of Human Genetics 73:1162-1169.
Stephens, M., N. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics 68:978-989.
Stephenson, J. R., and G. M. Calcarone. 1999. Southern California Mountains and Foothills Assessment: Habitat and Species Conservation Issues. Pp. 402. General Technical Report GTR-PSW-175. Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Albany, CA.
Szymura, J. M., and N. H. Barton. 1986. Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegata*, near Cracow in southern Poland. Evolution 40:1141-1159.
Szymura, J. M., and N. H. Barton. 1991. The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: Comparisons between transects and between loci. Evolution 45:237-261.
Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105.
Tamura, K., and M. Nei. 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.
Thomson, R. C., A. M. Shedlock, S. V. Edwards, and H. B. Shaffer. 2008. Developing markers for multilocus phylogenetics in non-model organisms: A test case with turtles. Molecular Phylogenetics and Evolution 49:514-525.
Ticehurst, C. B. 1938. A Systematic Review of the Genus *Phylloscopus*. Trustees of the British Museum, London.
Turelli, M., and L. C. Moyle. 2007. Asymmetric Postmating Isolation: Darwin's Corollary to Haldane's Rule. Genetics 176:1059-1088.
Vähä, J.-P., and C. R. Primmer. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology 15:63-72.
Van Devender, T. R. 1990. Late Quaternary vegetation and climate of the Sonoran Desert, United States and Mexico. Pp. 134-165 *in* J. L. Betancourt, T. R. Van Devender, and P. S. Martin, eds. Packrat Middens: Late Quaternary Environments of the Arid West. University of Arizona Press, Tucson.

Wake, D. B. 1997. Incipient species formation in salamanders of the *Ensatina* complex. Proceedings of the National Academy of Sciences 94:7761-7767.
Wake, D. B., and C. Schneider. 1998. Taxonomy of the plethodontid salamander genus *Ensatina*. Herpetologica 54:279-298.
Wake, D. B., and K. P. Yanev. 1986. Geographic variation in allozymes in a "ring species", the plethodontid salamander *Ensatina eschscholtzii* of western North America. Evolution 40:702-715.
Wake, D. B., K. P. Yanev, and C. W. Brown. 1986. Intraspecific sympatry in a ring species, the plethodontid salamander *Ensatina eschscholtzii*, in southern California. Evolution 40:866-868.
Wake, D. B., K. P. Yanev, and M. M. Frelow. 1989. Sympatry and hybridization in a "ring species": the Plethodontid salamander *Ensatina eschscholtzii*. Pp. 134-157 *in* D. Otte, and J. A. Endler, eds. Speciation and its Consequences. Sinauer Associates, Inc., Sunderland, MA.
Wakeley, J. 2008. Coalescent Theory: An Introduction. Roberts & Company Publishers.
Waltari, E., R. J. Hijmans, A. T. Peterson, A. S. Nyári, S. L. Perkins, and R. P. Guralnick. 2007. Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. PLoS ONE 2:e563.
Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358-1370.
Wieczorek, J., Q. Guo, and R. J. Hijmans. 2004. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. International Journal of Geographical Information Science 18:745-767.
Wirtz, P. 1999. Mother species-father species: unidirectional hybridization in animals with female choice. Animal Behaviour 58:1-12.
Wolfram, S. 1992. Redwood City: Addison-Wesley Publishing Company, Inc.
Wood, W. F. 1940. A new race of salamander, *Ensatina eschscholtzii picta*, from northern California and southern Oregon. University of California Publications in Zoology 42:425-428.

Appendix 1. Sampling localities for the allozyme dataset.

Population	Catalog No.	Locality	Latitude	Longitude
Crystal Creek	MVZ 185820	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185821	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185822	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185824	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185825	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185826	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185827	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185828	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185829	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
	MVZ 185830	Crystal Creek above Lucerne Valley, N side San Bernardino Mts.	34.37198	-116.92859
Sawmill Canyon	MVZ 185806	junction of W and N branches of Sawmill Canyon	34.05683	-116.85154
	MVZ 185848	junction of W and N branches of Sawmill Canyon	34.05683	-116.85154
	MVZ 185849	junction of W and N branches of Sawmill Canyon	34.05683	-116.85154
Fuller Mill Creek	MVZ 172528	Black Canyon, Upper Reaches of N Fork San Jacinto River above Hwy. 243	33.79326	-116.75203
	MVZ 172529	Black Canyon, Upper Reaches of N Fork San Jacinto River above Hwy. 243	33.79326	-116.75203
	MVZ 172530	Fuller Mill Creek vicinity below Camp Lawlor along Hwy. 243	33.79538	-116.74881
	MVZ 172531	Fuller Mill Creek vicinity below Camp Lawlor along Hwy. 243	33.79538	-116.74881
	MVZ 172533	Fuller Mill Creek vicinity below Camp Lawlor along Hwy. 243	33.79538	-116.74881
	MVZ 181960	vicinity of Fuller Mill Creek along Hwy. 243, Station 9D	33.79538	-116.74881
	MVZ 181938	Dark Canyon along San Jacinto River above Fuller Mill Creek	33.80308	-116.73046
Queen Creek	MVZ 185831	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185832	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185833	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185834	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185835	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185836	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185837	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185838	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185839	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
	MVZ 185840	Queen Creek 4.9 mi SE of Hwy. 74 on Santa Rosa Mt. Rd.	33.54286	-116.49009
Palomar	MVZ 181989	Lower Pine Valley, W Central border Palomar region, Site 3	33.30000	-116.80000
	MVZ 181990	Lower Pine Valley, W Central border Palomar region, Site 3	33.30000	-116.80000
	MVZ 181991	Lower Pine Valley, W Central border Palomar region, Site 3	33.30000	-116.80000
	MVZ 185808	Pine Valley, Palomar region, Station 4	33.31368	-116.80958
	MVZ 172596	Cedar Creek, 2.6 mi SE Palomar junction	33.27000	-116.85000
	MVZ 172597	Cedar Creek, 2.6 mi SE Palomar junction	33.27000	-116.85000

	MVZ 172598	Cedar Creek, 2.6 mi SE Palomar junction	33.27000	-116.85000
	MVZ 172599	Cedar Creek, 2.6 mi SE Palomar junction	33.27000	-116.85000
	MVZ 172600	Cedar Creek, 2.6 mi SE Palomar junction	33.27000	-116.85000
	MVZ 172601	Cedar Creek, 2.6 mi SE Palomar junction	33.27000	-116.85000
	MVZ 178750	E side Cedar Creek, above Hwy. S-7, Jeff Valley	33.30000	-116.83000
	MVZ 178751	E side Cedar Creek, above Hwy. S-7, Jeff Valley	33.30000	-116.83000
	MVZ 181405	Cedar Creek, at Jeff Valley, Palomar Mt.	33.29000	-116.83000
	MVZ 181974	Cedar Creek, at Jeff Valley, Palomar Mt.	33.29000	-116.83000
	MVZ 181975	Cedar Creek, at Jeff Valley, Palomar Mt.	33.29000	-116.83000
Cuyamaca	MVZ 167951	William Heise County Park, 4.6 mi SW Hwy. 78/79, W of Julian	33.03000	-116.59000
	MVZ 167952	William Heise County Park, 4.6 mi SW Hwy. 78/79, W of Julian	33.03000	-116.59000
	MVZ 167953	William Heise County Park, 4.6 mi SW Hwy. 78/79, W of Julian	33.03000	-116.59000
	MVZ 167954	William Heise County Park, 4.6 mi SW Hwy. 78/79, W of Julian	33.03000	-116.59000
	MVZ 167955	William Heise County Park, 4.6 mi SW Hwy. 78/79, W of Julian	33.03000	-116.59000
	MVZ 167956	William Heise County Park, 4.6 mi SW Hwy. 78/79, W of Julian	33.03000	-116.59000
	MVZ 167957	William Heise County Park, 4.6 mi SW Hwy. 78/79, W of Julian	33.03000	-116.59000
	MVZ 167958	William Heise County Park, 4.6 mi SW Hwy. 78/79, W of Julian	33.03000	-116.59000
	MVZ 169051	W of Julian, 4.6 mi SW Hwy. 78 and Hwy. 79, William Heise County Park	33.03000	-116.59000
	MVZ 181392	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181393	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181394	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181395	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181396	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181397	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181398	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181399	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181400	Juch Canyon, 1.2 mi E Hwy. 78 and Hwy. 79 on Wynola Rd.	33.09755	-116.62454
	MVZ 181327	Camp Wolahi	32.99000	-116.59000
	MVZ 181401	E Colby Spring, Camp Wolahi, Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181402	E Colby Spring, Camp Wolahi, Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181403	E Colby Spring, Camp Wolahi, Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181404	Camp Wolahi, S side of Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181941	Camp Wolahi, S side of Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181942	Camp Wolahi, S side of Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181944	above Pichacho Trail, E Colby Spring Camp Wolahi, S side Boulder Creek Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181945	above Pichacho Trail, E Colby Spring Camp Wolahi, S side Boulder Creek Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181946	above Pichacho Trail, E Colby Spring Camp Wolahi, S side Boulder Creek Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181947	E Colby Spring above Trail Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
	MVZ 181948	E Colby Spring above Trail Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000

MVZ 181949	E Colby Spring above Trail Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181950	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181951	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181952	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181953	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181954	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181955	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181956	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181957	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181958	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181959	by outhouse below trail and E Colby Spring, Camp Wolahi, S side Boulder Creek, Cuyamaca Mts.	32.99000	-116.59000
MVZ 181967	W side Colby Spring, Camp Wolahi, side Boulder Creek, Cuyamaca Mts., Stations3B and 4B	32.99000	-116.59000
MVZ 181968	W side Colby Spring, Camp Wolahi, side Boulder Creek, Cuyamaca Mts., Stations3B and 4B	32.99000	-116.59000
MVZ 181969	W side Colby Spring, Camp Wolahi, side Boulder Creek, Cuyamaca Mts., Stations3B and 4B	32.99000	-116.59000
MVZ 181970	W side Colby Spring, Camp Wolahi, side Boulder Creek, Cuyamaca Mts., Stations3B and 4B	32.99000	-116.59000
MVZ 181971	W side Colby Spring, Camp Wolahi, side Boulder Creek, Cuyamaca Mts., Stations3B and 4B	32.99000	-116.59000
MVZ 181415	Engineers Rd., 3.5 mi W (by road) Hwy. 79	33.01000	-116.61000
MVZ 181416	Engineers Rd., 3.5 mi W (by road) Hwy. 79	33.01000	-116.61000
MVZ 181417	Engineers Rd., 3.5 mi W (by road) Hwy. 79	33.01000	-116.61000
MVZ 181418	Engineers Rd., 3.5 mi W (by road) Hwy. 79	33.01000	-116.61000
MVZ 181419	Engineers Rd., 3.5 mi W (by road) Hwy. 79	33.01000	-116.61000
MVZ 181420	Engineers Rd., 3.5 mi W (by road) Hwy. 79	33.01000	-116.61000

Appendix 2. Sampling for the mitochondrial DNA and microsatellite datasets.

Population	Tissue or field number	Locality	Latitude	Longitude	mtDNA	msats
Crystal Creek	JFP 517	Crystal Creek, San Bernardino Mountains	34.37198	-116.92859	x	x
	TJD 273	Arctic Canyon, N slope of San Bernardino Mts.	34.34330	-116.89481	x	x
	TJD 300	Marble Canyon, San Bernardino Mts.	34.33391	-116.86587	x	x
Fuller Mill Creek	TJD 432	Dark Canyon, S of campground, San Jacinto Mts.	33.80206	-116.73317	x	x
	TJD 433	Dark Canyon, S of campground, San Jacinto Mts.	33.80206	-116.73317	x	x
	TJD 430	Dark Canyon, S of campground, San Jacinto Mts.	33.80179	-116.73401	x	x
	TJD 431	Dark Canyon, S of campground, San Jacinto Mts.	33.80153	-116.73379	x	x
	TJP 29472A	Fuller Mill Creek Picnic Area, San Jacinto Mts.	33.79983	-116.73250	x	x
	TJD 435	Dark Canyon, San Jacinto Mts.	33.79966	-116.73253		x
	TJD 436	Dark Canyon, San Jacinto Mts.	33.79966	-116.73253	x	x
	TJD 437	Dark Canyon, San Jacinto Mts.	33.79951	-116.73322	x	x
	TJD 434	Dark Canyon, N of Hwy. 243, San Jacinto Mts.	33.79929	-116.73397	x	x
Queen Creek	TJD 438	Queen Creek, 4.6 mi. (rd.) SE of Hwy. 74 on FS rd. 7S02, Santa Rosa Mts.	33.54291	-116.48181	x	x
	TJD 439	Queen Creek, 4.6 mi. (rd.) SE of Hwy. 74 on FS rd. 7S02, Santa Rosa Mts.	33.54274	-116.48120	x	x
Palomar	TJD 495	vicinity of observatory, Palomar Mountain	33.35368	-116.86920	x	x
	TJD 503	Fry Creek Campground, Palomar Mountain	33.34402	-116.88315	x	x
	TJD 358	btwn. Thunder Spring & Scott's Cabin trails, Palomar Mtn. SP	33.33799	-116.90012	x	x
	TJD 365	above Thunder Springs Trail	33.33793	-116.90147	x	x
	TJD 380	above Thunder Springs Trail	33.33774	-116.90147	x	x
	TJD 360	btwn. Thunder Spring & Scott's Cabin trails, Palomar Mtn. SP	33.33748	-116.89946		x
	TJD 359	btwn. Thunder Spring & Scott's Cabin trails, Palomar Mtn. SP	33.33735	-116.90030	x	x
	TJD 372	Scott's Cabin Trail, Palomar Mtn. SP	33.33446	-116.90500	x	x
	TJD 369	Scott's Cabin Trail, Palomar Mtn. SP	33.33414	-116.90368	x	x
	TJD 377	above Thunder Springs Trail	33.33712	-116.89956	x	x
	TJD 363	above Thunder Springs Trail	33.33679	-116.90053	x	x
	TJD 489	side of State Park Road between employee residences	33.33646	-116.90789		x
	TJD 491	side of State Park Road between employee residences	33.33619	-116.90780	x	x
	TJD 502	behind general store at jct. S6/S7	33.31435	-116.86621	x	x
Hot Springs	MVZ 207642	road to Lookout, 4.3 mi N Camino San Ignacio, Hot Springs Mt.	33.30814	-116.55264	x	x
	MVZ 207643	road to Lookout, 4.3 mi N Camino San Ignacio, Hot Springs Mt.	33.30814	-116.55264	x	
	MVZ 236191	Along Agua Caliente Creek near 32994 Camino Moro, Los Tules	33.28633	-116.61800	x	
	TJD 464	Hot Springs Mtn., 7.6 mi. (rd.) E of Hwy. 79 on road to lookout	33.27663	-116.53803	x	x
	TJD 465	Hot Springs Mtn., 7.6 mi. (rd.) E of Hwy. 79 on road to lookout	33.27663	-116.53803	x	x
	TJD 466	Hot Springs Mtn., 7.6 mi. (rd.) E of Hwy. 79 on road to lookout	33.27663	-116.53803	x	x
	TJD 467	Hot Springs Mtn., 7.6 mi. (rd.) E of Hwy. 79 on road to lookout	33.27663	-116.53803	x	x
Cuyamaca	TJD 459	Volcan Mountain	33.11913	-116.61294	x	x
	TJD 460	Volcan Mountain	33.11913	-116.61294	x	x
	TJD 461	Volcan Mountain	33.11913	-116.61294	x	x
	TJD 462	Volcan Mountain	33.11901	-116.61669	x	x
	TJD 463	Volcan Mountain	33.11739	-116.61454	x	x
	TJD 497	near camp below Milk Ranch Road, Cuyamaca Rancho State Park	32.97354	-116.58454	x	x

	TJD 496	near camp below Milk Ranch Road, Cuyamaca Rancho State Park	32.97312	-116.58501	x	x
	TJD 498	near camp below Milk Ranch Road, Cuyamaca Rancho State Park	32.97165	-116.58382	x	x
	TJD 499	near camp below Milk Ranch Road, Cuyamaca Rancho State Park	32.97165	-116.58382	x	x
	TJD 500	near camp below Milk Ranch Road, Cuyamaca Rancho State Park	32.97165	-116.58382	x	x
	TJD 501	near camp below Milk Ranch Road, Cuyamaca Rancho State Park	32.97165	-116.58382	x	x
	TJD 468	near camp below Milk Ranch Road, Cuyamaca Rancho State Park	32.97156	-116.58320	x	x
	TJD 469	near camp below Milk Ranch Road, Cuyamaca Rancho State Park	32.97086	-116.58320	x	x
	TJD 447	Paso Picacho Campground, Cuyamaca Rancho State Park	32.96022	-116.58245	x	x
	TJD 445	Paso Picacho Campground, Cuyamaca Rancho State Park	32.96006	-116.58203	x	x
	TJD 446	Paso Picacho Campground, Cuyamaca Rancho State Park	32.95989	-116.58162	x	x
	TJD 444	Paso Picacho Campground, Cuyamaca Rancho State Park	32.95953	-116.58123	x	x
	TJD 453	Green Valley Campground, ranger's house, Cuyamaca Rancho State Park	32.90287	-116.50898	x	x
	TJD 451	Green Valley Campground, ranger's house, Cuyamaca Rancho State Park	32.90237	-116.58069	x	x
	TJD 452	Green Valley Campground, ranger's house, Cuyamaca Rancho State Park	32.90237	-116.58075	x	x
	TJD 448	Green Valley Campground, ranger's house, Cuyamaca Rancho State Park	32.90222	-116.58009	x	x
	TJD 449	Green Valley Campground, ranger's house, Cuyamaca Rancho State Park	32.90222	-116.58009	x	x
	TJD 450	Green Valley Campground, ranger's house, Cuyamaca Rancho State Park	32.90222	-116.58009	x	x
	TJD 458	Noble Canyon National Recreation Trail, Pine Valley	32.85190	-116.52313	x	x
	MVZ 195628	Mt. Laguna Rd. at Kitchen Creek Rd., Laguna Mts.	32.85000	-116.47000	x	x
	TJD 457	Kitchen Creek, just south of Sunrise Hwy. (S1)	32.84376	-116.44572	x	x
	TJD 454	Cibbetts Flat Campground, Kitchen Creek, Laguna Mts.	32.77759	-116.44776	x	x
	TJD 455	Cibbetts Flat Campground, Kitchen Creek, Laguna Mts.	32.77759	-116.44776	x	x
	TJD 456	Cibbetts Flat Campground, Kitchen Creek, Laguna Mts.	32.77759	-116.44776	x	x
Sierra de Juarez	SDFIELD 1158	2.0 km N of Rancho Bajia Largo del Sur, Sierra de Juarez	31.87173	-115.93084	x	x
SSPM	SDFIELD 1343	La Tasajera Region, Sierra San Pedro Mártir	30.99280	-115.50910	x	x
	SDFIELD 1167	La Tasajera Region, Sierra San Pedro Mártir	30.95832	-115.50283	x	x
	MVZ 229220	vicinity of La Tasajera, Sierra San Pedro Mártir	30.93333	-115.56667	x	x
San Quintín	SDFIELD 573	vic. San Quintín, Baja California	30.46853	-116.00833	x	x
	SDFIELD 574	vic. San Quintín, Baja California	30.46853	-116.00833	x	x
	SDFIELD 736	vic. San Quintín, Baja California	30.46853	-116.00833	x	x
	SDFIELD 901	vic. San Quintín, Baja California	30.46853	-116.00833	x	x

Appendix 3. Summary statistics for 19 allozyme loci the Large-blotched Ensatina. *N* = sample size, % miss. = percent missing data per locus, N_A = number of alleles per locus, number of expected (H_E) and observed (H_O) heterozygotes, F_{IS} = fixation index following the method of Weir and Cockerham (1984), All = global F_{IS} calculated using Fisher's method. Expected number of heterozygotes computed using Levene's (1949) correction. *P*-values in bold are significant.

		Locus																	
Population		Loc2	Loc3	Loc6	Loc7	Loc8	Loc10	Loc11	Loc12	Loc13	Loc14	Loc15	Loc16	Loc17	Loc18	Loc19	Loc20	Loc21	Loc23
Crystal Creek	*N*	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	% miss.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N_A	1	1	1	1	1	1	1	1	1	2	1	1	2	2	0	2	1	1
	H_E	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.05	0.00	0.00	1.00	5.21	0.00	5.21	0.00	0.00
	H_O	0	0	0	0	0	0	0	0	0	4	0	0	1	5	0	5	0	0
	F_{IS}	---	---	---	---	---	---	---	---	---	0.22	---	---	---	0.04	---	0.04	---	---
	P	---	---	---	---	---	---	---	---	---	0.57	---	---	---	1.00	---	1.00	---	---
Sawmill Canyon	*N*	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	% miss.	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N_A	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	2	2	1
	H_E	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.80	0.00	1.60	0.00	1.80	1.60	0.00
	H_O	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	1	2	0
	F_{IS}	---	---	---	---	---	---	---	---	---	---	---	-1.00	---	1.00	---	0.50	-0.33	---
	P	---	---	---	---	---	---	---	---	---	---	---	0.40	---	0.20	---	1.00	1.00	---
Fuller Mill Creek	*N*	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
	% miss.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N_A	1	2	1	1	2	1	1	2	2	1	1	1	1	3	2	2	2	2
	H_E	0.00	3.46	0.00	0.00	1.00	0.00	0.00	3.69	1.00	0.00	0.00	0.00	0.00	4.54	1.00	3.77	3.69	2.54
	H_O	0	1	0	0	1	0	0	4	1	0	0	0	0	4	1	1	2	3
	F_{IS}	---	0.73	---	---	---	---	---	-0.09	---	---	---	---	---	0.13	---	0.75	0.48	-0.20
	P	---	0.11	---	---	---	---	---	1.00	---	---	---	---	---	1.00	---	0.12	0.44	1.00
Queen Creek	*N*	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	% miss.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N_A	2	2	1	1	1	1	1	3	2	2	1	2	1	2	1	2	2	2
	H_E	2.68	2.68	0.00	0.00	0.00	0.00	0.00	5.63	2.68	1.89	0.00	1.89	0.00	4.79	0.00	5.21	1.89	5.26
	H_O	3	3	0	0	0	0	0	8	3	2	0	2	0	7	0	3	2	6
	F_{IS}	-0.13	-0.13	---	---	---	---	---	-0.45	-0.13	-0.06	---	-0.06	---	-0.50	---	0.44	-0.06	-0.15
	P	1.00	1.00	---	---	---	---	---	0.11	1.00	1.00	---	1.00	---	0.22	---	0.25	1.00	1.00
Palomar	*N*	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
	% miss.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N_A	2	1	1	1	3	1	1	2	1	1	2	2	2	3	2	1	1	1
	H_E	1.00	0.00	0.00	0.00	5.24	0.00	0.00	6.07	0.00	0.00	7.72	3.59	7.21	4.45	7.21	0.00	0.00	0.00
	H_O	1	0	0	0	4	0	0	4	0	0	6	4	3	5	5	0	0	0
	F_{IS}	---	---	---	---	0.24	---	---	0.35	---	---	0.23	-0.12	0.59	-0.13	0.31	---	---	---
	P	---	---	---	---	**0.05**	---	---	0.23	---	---	0.61	1.00	**0.04**	1.00	0.29	---	---	---
Cuyamaca	*N*	52	52	52	52	52	52	52		52	52		52	52	52		52	52	
	% miss.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.00	0.06	0.00	0.00	0.02

	N_A	1	2	2	2	2	2	2	2	2	2	3	2	3	3	2	1	4	2
	H_E	0.00	1.98	1.98	1.98	9.93	13.67	4.81	25.71	1.98	2.94	23.46	26.10	17.29	22.95	16.67	0.00	15.10	7.45
	H_O	0	2	2	2	9	8	5	29	2	3	26	18	17	22	15	0	13	6
	F_{IS}	---	-0.01	---	-0.01	0.09	0.42	-0.04	-0.13	-0.01	-0.02	-0.11	0.31	0.02	0.04	0.10	---	0.14	0.20
	P	---	1.00	---	1.00	0.44	**0.01**	1.00	0.41	1.00	1.00	0.77	**0.03**	**0.01**	0.20	0.67	---	0.25	0.26
All (Fisher's)	P	---	0.62	---	---	0.11	---	---	0.33	1.00	0.98	0.82	0.35	**0.00**	0.66	0.51	0.53	0.81	0.84
	% miss.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.03	0.00	0.00	0.01

Appendix 4. Summary statistics for 10 microsatellite loci for Large-blotched Ensatina populations. N = sample size, % missing = percent missing data per locus, N_A = number of alleles per locus, number of expected (H_E) and observed (H_O) heterozygotes, F_{IS}=fixation index following the method of Weir and Cockerham (1984). Expected number of heterozygotes computed using Levene's (1949) correction. *P*-values in bold are significant. Maximum-likelihood null allele frequencies were estimated using the EM algorithm (Dempster et al. 1977).

		Locus									
Population		ENS1	ENS3	ENS4	ENS5	ENS6	ENS7	ENS11	ENS13	ENS15	ENS20
Crystal Creek	*N*	3	1	0	0	3	3	1	3	3	3
	% missing	0.00	0.67	1.00	1.00	0.00	0.00	0.67	0.00	0.00	0.00
	N_A	3	2	---	---	3	4	2	4	5	3
	H_E	1.800	---	---	---	2.200	2.400	2.600	2.600	2.800	2.200
	H_O	2	---	---	---	2	2	3	3	3	2
	F_{IS}	-0.143	---	---	---	0.111	0.200	---	-0.200	-0.091	0.111
	P	1.000	---	---	---	1.000	0.603	---	1.000	1.000	1.000
	Null freq.	0.000	---	---	---	0.000	0.000	---	0.000	0.000	0.000
Fuller Mill Creek	*N*	9	9	9	9	9	9	7	9	8	9
	% missing	0.00	0.00	0.00	0.00	0.00	0.00	0.78	0.00	0.11	0.00
	N_A	9	6	6	9	9	10	2	7	9	9
	H_E	8.177	7.471	7.706	7.588	8.059	8.412	1.000	7.588	7.400	8.177
	H_O	9	7	7	7	9	9	1	4	8	8
	F_{IS}	-0.108	0.067	0.097	0.082	-0.125	-0.075	---	0.488	-0.087	0.023
	P	0.811	0.970	0.096	0.436	0.900	1.000	---	**0.021**	1.000	0.396
	Null freq.	0.000	0.000	0.021	0.000	0.000	0.000	---	0.198	0.000	0.000
Queen Creek	*N*	2	2	1	2	2	2	2	2	2	2
	% missing	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N_A	2	4	2	3	2	4	3	2	3	4
	H_E	1.333	2.000	---	1.667	1.000	2.000	1.667	1.000	1.667	2.000
	H_O	0	2	---	1	1	2	2	1	2	2
	F_{IS}	1.000	0.000	---	0.500	---	0.000	-0.333	---	-0.333	0.000
	P	0.331	1.000	---	0.331	---	1.000	1.000	---	1.000	1.000
	Null freq.	---	0.000	---	0.000	---	0.000	0.000	---	0.000	0.000
Palomar	*N*	13	14	9	14	14	14	13	4	5	13
	% missing	0.07	0.00	0.36	0.00	0.00	0.00	0.07	0.43	0.71	0.07
	N_A	8	11	5	10	10	10	10	4	4	16
	H_E	10.600	7.353	7.353	12.037	12.222	12.444	11.560	5.800	3.286	12.480
	H_O	7	2	2	12	10	13	11	5	1	13
	F_{IS}	0.349	-0.090	0.740	0.003	0.188	-0.046	0.050	0.146	0.727	-0.044
	P	**0.009**	1.000	**0.001**	0.721	0.132	0.948	**0.044**	0.504	**0.031**	1.000
	Null freq.	**0.166**	0.000	**0.315**	0.000	0.066	0.000	0.033	0.053	0.278	0.000
Hot Springs	*N*	5	5	3	5	5	4	0	1	5	5
	% missing	0.00	0.00	0.40	0.00	0.00	0.20	1.00	0.80	0.00	0.00
	N_A	4	6	2	8	6	5	---	1	6	7
	H_E	4.000	4.556	1.000	4.778	4.333	3.571	---	---	4.222	4.556
	H_O	2	5	1	4	5	4	---	---	3	5
	F_{IS}	0.529	-0.111	---	0.180	-0.177	-0.143	---	---	0.314	-0.111
	P	0.145	1.000	---	0.270	**0.015**	1.000	---	---	0.175	1.000
	Null freq.	0.180	0.000	---	0.133	0.000	0.000	---	---	0.086	0.000
Cuyamaca	*N*	29	29	26	28	27	28	27	20	24	27
	% missing	0.00	0.00	0.10	0.03	0.07	0.03	0.07	0.24	0.17	0.07
	N_A	10	13	11	10	13	18	13	10	18	15
	H_E	25.772	24.860	23.471	24.073	23.660	26.018	23.830	19.023	21.702	24.321
	H_O	15	23	14	24	22	25	20	11	13	21
	F_{IS}	0.422	0.076	0.408	0.003	0.071	0.040	0.163	0.428	0.406	0.139
	P	**0.000**	0.297	**0.000**	0.037	0.216	0.381	**0.012**	**0.000**	**0.000**	0.223
	Null freq.	0.192	0.042	0.189	0.000	0.040	0.016	0.085	0.183	0.176	0.042
Sierra de Juarez	*N*	1	1	1	1	1	1	1	1	1	1
	% missing	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N_A	2	2	1	2	2	1	2	2	2	1

	H_E	---	---	---	---	---	---	---	---	---	---
	H_O	---	---	---	---	---	---	---	---	---	---
	F_{IS}	---	---	---	---	---	---	---	---	---	---
	P	---	---	---	---	---	---	---	---	---	---
	Null freq.	---	---	---	---	---	---	---	---	---	---
Sierra San Pedro Mártir	*N*	3	1	1	3	3	3	2	3	3	2
	% missing	0.00	0.67	0.67	0.00	0.00	0.00	0.33	0.00	0.00	0.33
	N_A	3	1	2	4	5	3	3	6	6	2
	H_E	2.200	---	---	2.600	2.800	2.400	1.667	3.000	3.000	1.333
	H_O	2	---	---	3	3	2	1	3	3	0
	F_{IS}	0.111	---	---	0.200	-0.091	0.200	0.500	0.000	0.000	1.000
	P	1.000	---	---	1.000	1.000	0.468	0.329	1.000	1.000	0.335
	Null freq.	0.000	---	---	0.000	0.000	0.000	0.000	0.000	0.000	---
San Quintín	*N*	4	4	0	0	4	4	3	4	4	4
	% missing	0.00	0.00	1.00	1.00	0.00	0.00	0.25	0.00	0.00	0.00
	N_A	2	3	---	---	1	2	1	2	3	2
	H_E	1.714	2.857	---	---	0.000	1.714	0.000	1.714	2.429	1.714
	H_O	2	4	---	---	0	2	0	0	3	2
	F_{IS}	0.111	-0.500	---	---	---	-0.200	---	1.000	-0.286	-0.200
	P	1.000	0.548	---	---	---	1.000	---	0.147	1.000	1.000
	Null freq.	---	0.000	---	---	---	---	---	---	0.000	---
All (Fisher's method)	*P*	**0.00**	0.99	**0.00**	0.32	0.20	1.00	**0.03**	**0.00**	**0.00**	0.97
	% missing	0.01	0.06	0.29	0.11	0.03	0.03	0.27	0.24	0.23	0.06